INTERVAL LINEAR SYSTEMS

by

J. Rohn, Prague

0. Introduction and notations

This paper brings miscellaneous results concerning interval linear systems, grouped into relatively independent sections and covering such areas as properties of the solution set, computational aspects, regularity, nonnegative invertibility and interval linear programming. The reader is strongly recommended to read first our paper [22] to which references are frequently made in the text.

Notations used are almost the same as in [22]. We repeat them briefly. We deal here with interval linear systems $A^I x = b^I$, where $A^I = [A_c - \Delta, A_c + \Delta] = [\underline{A}, \overline{A}]$ is an nxn interval matrix, $b^I = [b_c - \delta, b_c + \delta] = [\underline{b}, \overline{b}]$ is an interval n-vector, $n \geq 1$. We denote $N = \{1,\dots,n\}$, $e = (1,\dots,1)^T$, $f = -e$, $Y = \{y \in R^n; |y| = e\}$; T_t is the diagonal matrix with diagonal $t \in R^n$. For each $t, z \in R^n$, we define $A_{tz} = A_c - T_t \Delta T_z$, $b_t = b_c + T_t \delta$ (i.e. $(A_{tz})_{ij} = (A_c)_{ij} - t_i \Delta_{ij} z_j$, $(b_t)_i = (b_c)_i + t_i \delta_i$, $i, j \in N$); if $|t| \leq e$ and $|z| \leq e$ (especially if $t \in Y$ and $z \in Y$), then $A_{tz} \in A^I$, $b_t \in b^I$. For each $x \in R^n$, we define $\operatorname{sgn} x \in Y$ by $(\operatorname{sgn} x)_i = 1$ if $x_i \geq 0$ and $(\operatorname{sgn} x)_i = -1$ if $x_i < 0$. $x \in M$ is called an extremal point of a convex set M if there does not exist a pair $x_1, x_2 \in M$ such that $x_1 \neq x_2$ and $x = \frac{1}{2}(x_1 + x_2)$. $\rho(A)$ denotes the spectral radius of A.

AMS Subject Classifications: 65G10, 90C05, 90C33

1. The solution set

In this section, we present some properties of the solution set $X = \{x;\ Ax = b,\ A \in A^I,\ b \in b^I\}$. Some of them are known; we give new proofs here, sometimes shorter.

The basic description of the solution set X is due to Oettli and Prager [16] :

Theorem 1.1. $X = \{x;\ |A_c x - b_c| \le \Delta|x| + \delta\}$.

Proof. If $Ax = b$ for some $A \in A^I$, $b \in b^I$, then $|A_c x - b_c| = |(A_c - A)x + b - b_c| \le \Delta|x| + \delta$. Conversely, let the inequality hold; define $t \in R^n$ by

$$t_i = \begin{cases} (A_c x - b_c)_i / (\Delta|x| + \delta)_i & \text{if } (\Delta|x| + \delta)_i \ne 0 \\ 1 & \text{otherwise} \end{cases}$$

$(i \in N)$, then $A_c x - b_c = T_t(\Delta|x| + \delta)$ and with $z = \operatorname{sgn} x$ we obtain $A_{tz} x = b_t$, which, since $t \in [f,e]$, implies $x \in X$. ■

We have simultaneously proved this

Corollary 1.1. Let $x \in X$. Then it satisfies $A_{tz} x = b_t$ for $z = \operatorname{sgn} x$ and some $t \in [f,e]$.

The following two results are due to Beeck [4] :

Corollary 1.2. $X = \{x;\ A_{ez} x \le \overline{b},\ A_{fz} x \ge \underline{b},\ T_z x \ge 0,\ z \in Y\}$.

Proof. Follows immediately from theorem 1.1 when expressing $|x|$ by $T_z x$ with $z = \operatorname{sgn} x$. ■

Corollary 1.3. The intersection of X with each orthant is a convex polytope.

Proof. From corollary 1.2, with z fixed.■

A still another description of X can be obtained from theorem 1.1 when substituting $x = x^+ - x^-$, $|x| = x^+ + x^-$:

<u>Corollary 1.4.</u> $X = \{x;\ \underline{A}x^+ - \overline{A}x^- \leq \overline{b},\ \overline{A}x^+ - \underline{A}x^- \geq \underline{b}\}$.

The central result concerning the solution set is theorem 2 in [22]. We give here another proof, not using theorem 0, and perhaps more insightful. We begin with this theorem:

<u>Theorem 1.2.</u> Let A^I be regular. Then each extremal point x of Conv X satisfies $|A_c x - b_c| = \Delta|x| + \delta$.

<u>Proof.</u> x, being an extremal point of Conv X, must lie in X. Hence $A_{tz}x = b_t$ for $z = \operatorname{sgn} x$ and $t \in [f,e]$ (corollary 1.1). Assume for contrary that $t_j \in (-1,1)$ for some j. Then there exists a sufficiently small $\varepsilon > 0$ such that for each $t' = t + \gamma e_j$, $\gamma \in (-\varepsilon, \varepsilon)$, the solution x' of $A_{t'z}x' = b_{t'}$ is, according to lemma 1 in [22], given by $x' = x - \varphi(\gamma)A_{tz}^{-1}e_j$, where

$$\varphi(\gamma) = \frac{\gamma(\Delta|x| + \delta)_j}{1 - \gamma(\Delta T_z A_{tz}^{-1})_{jj}} .$$

Hence x is an interior point of a segment lying in X, so that it cannot be an extremal point of Conv X. Thus $|t| = e$; since $A_{tz}x = b_t$ implies $A_c x - b_c = T_t(\Delta|x| + \delta)$, we obtain $|A_c x - b_c| = \Delta|x| + \delta$. ∎

As in [22], we can bring the equation $|A_c x - b_c| = \Delta|x| + \delta$ by substituting $y = \operatorname{sgn}(A_c x - b_c)$, $x = x^+ - x^-$, $|x| = x^+ + x^-$ to the form

$$A_{ye}x^+ - A_{yf}x^- = b_y \tag{1.1}$$

$(y \in Y)$.

<u>Theorem 1.3.</u> Let A^I be regular. Then for each $y \in Y$, the equation (1.1) has exactly one solution $x_y \in X$ and we have Conv X = $= \operatorname{Conv}\{x_y;\ y \in Y\}$.

Proof. As in [22], the existence and unicity of the x_y's follows from the P-property of matrices $A_{ye}^{-1}A_{yf}$ $(y \in Y)$. According to theorem 1.2, Conv X = Conv $\{x;\ |A_c x - b_c| = \Delta|x| + \delta\} =$
$=$ Conv $\{x_y;\ y \in Y\}$. ∎

A system $A^I x = b^I$ is called normal if $\Delta|x| + \delta > 0$ for each $x \in X$. It can be easily shown that a system is normal if either of the following two conditions holds: (i) $\delta > 0$, (ii) $\Delta > 0$ and $0 \notin b^I$.

Theorem 1.4. Let $A^I x = b^I$ be a normal system with a regular matrix A^I. Then the mapping $x \to t$ given by

$$t_i = \frac{(A_c x - b_c)_i}{(\Delta|x| + \delta)_i} \qquad (i \in N) \qquad (1.2)$$

is a continuous one-to-one mapping of X onto the interval $[f,e]$.

Proof. It follows from theorem 1.1 that X is mapped into $[f,e]$. Let $t \in [f,e]$; then, according to theorem A in [23], the linear complementarity problem

$$x^+ = A_{te}^{-1}A_{tf}x^- + A_{te}^{-1}b_t \qquad (1.3)$$

has exactly one solution x. (1.3) can be rearranged to $A_c x - b_c = T_t(\Delta|x| + \delta)$, which shows that (1.2) holds, hence t is the image of x, and x is unique. ∎

The inverse mapping $t \to x$ cannot be expressed explicitly, but for each t the inverse image x may be computed by solving (1.3) using Murty's algorithm [13]. Notice that each $y \in Y$ is mapped exactly on the point x_y. 0 is mapped on $x_c = A_c^{-1}b_c$. For each $z \in Y$, the set $\{t;\ t \in [f,e],\ T_z t \geq 0\}$ is mapped onto the intersection of X with the cone emanating from x_c and spanned over the vectors $z_j(A_c^{-1})_{.j}$ $(j \in N)$.

Corollary 1.5. Let $A^I x = b^I$ be a normal system with a regular matrix A^I. Then for each $y, y' \in Y$, $y \neq y'$ implies $x_y \neq x_{y'}$.

Proof. From theorem 1.4.■

2. Nonconvexity of the solution set

It follows from corollary 1.3 that the solution set X, being a union of 2^n convex polytopes, is generally nonconvex. In this section we shall examine the problem of (non)convexity in terms of the x_y's.

Theorem 2.1. Let A^I be regular. Then X is nonconvex if and only if there exist $y, y' \in Y$ and $i, j \in N$ such that $y_i y'_i = 1$, $\Delta_{ij} > 0$ and $(x_y)_j (x_{y'})_j < 0$.

Proof. "If": Let $x = \lambda x_y + \mu x_{y'}$ for some $\lambda > 0$, $\mu > 0$, $\lambda + \mu = 1$. Then $|x|_j < \lambda |x_y|_j + \mu |x_{y'}|_j$, hence $|A_c x - b_c|_i = |y_i (\Delta (\lambda |x_y| + \mu |x_{y'}|) + \delta)_i| > (\Delta |x| + \delta)_i$, showing that $x \notin X$ (theorem 1.1). Thus no interior point of the segment connecting x_y with $x_{y'}$ belongs to X, so X is nonconvex. "Only if": Assume for contrary that for each $y, y' \in Y$ and $i, j \in N$, $y_i y'_i = 1$ and $\Delta_{ij} > 0$ imply $(x_y)_j (x_{y'})_j \geq 0$. Let $x = \sum_Y \lambda_y x_y$, $\lambda_y \geq 0$ for each $y \in Y$, $\sum_Y \lambda_y = 1$. Then for each $i \in N$ we have $(A_c x - b_c)_i =$
$$= \sum_Y \lambda_y y_i (\sum_j \Delta_{ij} |x_y|_j + \delta_i) = \sum_j \Delta_{ij} (|\sum_{y_i = 1} \lambda_y x_y|_j - |\sum_{y_i = -1} \lambda_y x_y|_j) + \sum_Y \lambda_y y_i \delta_i,$$
hence $|A_c x - b_c|_i \leq \sum_j \Delta_{ij} |\sum_Y \lambda_y x_y|_j + \delta_i = (\Delta |x| + \delta)_i$, so that $x \in X$. Thus we have shown that $\mathrm{Conv}\{x_y;\ y \in Y\} \subset X$, which in view of theorem 1.3 means that Conv X = X, and X is convex.■

Together with this result, we have proved

Corollary 2.1. Let A^I be regular. Then, X is nonconvex if and only if there exist $x_y, x_{y'}$, $x_y \neq x_{y'}$, such that no interior point of the segment connecting x_y with $x_{y'}$ belongs to X.

Theorem 2.2. Let A^I be regular and let $\Delta > 0$. Then X is nonconvex if and only if there exist $y, y' \in Y$, $y \neq -y'$, such that $(x_y)_j (x_{y'})_j < 0$ for some $j \in N$.

The proof follows immediately from theorem 2.1. The reason for the assumption $y \neq -y'$ is explained in the next theorem:

Theorem 2.3. Let A^I be regular. Then for each $y \in Y$, the whole segment connecting x_y with x_{-y} lies in X.

Proof. Take $x = \lambda x_y + \mu x_{-y}$, $\lambda > 0$, $\mu > 0$, $\lambda + \mu = 1$. Then from $A_c x_y - b_c = T_y(\Delta|x_y| + \delta)$, $A_c x_{-y} - b_c = -T_y(\Delta|x_{-y}| + \delta)$ we get $|A_c x - b_c| = |T_y(\Delta(|\lambda x_y| - |\mu x_{-y}|) + (\lambda - \mu)\delta)| \leq$ $\leq \Delta|x| + \delta$, hence $x \in X$. ∎

The solution set can be convex even if it intersects the interiors of all orthants:

Example 2.1. The solution set of the system

$$[1,2]x_1 + [1,2]x_2 = [-1,1]$$
$$[1,2]x_1 + [-2,-1]x_2 = [-1,1]$$

is a square with vertices $(1,0)^T$, $(0,1)^T$, $(-1,0)^T$, $(0,-1)^T$.

The following theorem will be used in interval linear programming (section 10):

Theorem 2.4. Let $A^I x = b^I$ be a normal system with a regular matrix A^I such that $x_y > 0$ for each $y \in Y$. Then X is a convex polyhedron with vertices x_y $(y \in Y)$ and two vertices $x_y, x_{y'}$ are neighbouring if and only if y and y' differ in just one entry.

Proof. It follows from theorem 1.3 and corollary 1.2 that $X = \{ x;\ \underline{A}x \leq \overline{b},\ \overline{A}x \geq \underline{b},\ x \geq 0\}$. Each extremal point (vertex) of X is equal to some x_y according to theorem 1.2. The fact that each x_y is an extremal point of X follows from theorem 2 in [20]. If x_y and $x_{y'}$ are neighbouring, then the edge connecting them must lie in the intersection of n-1 hyperplanes. Since the normality precludes the possibility of $(\underline{A}x)_i = \overline{b}_i, (\overline{A}x)_i = \underline{b}_i$ holding simultaneously for some i, there must exist a $\tilde{y} \in Y$ and a $k \in N$ such that each point of the edge satisfies the system

$$(A_{\tilde{y}e}x)_i = (b_{\tilde{y}})_i, \qquad i \in N - \{k\} . \tag{2.1}$$

Hence x_y and $x_{y'}$ also satisfy (2.1), so that y and y' differ only in the k-th entry. ∎

3. Computational complexity of the sign accord algorithm

In [22], we proposed using Murty's algorithm [13] for solving the linear complementarity problem $x^+ = A_{ye}^{-1}A_{yf}x^- + A_{ye}^{-1}b_y$. This requires inverting A_{ye} and then working with a simplex-like tableau. We also showed in [22] that Murty's algorithm can be reformulated in terms of solving systems $A_{yz}x = b_y$ until $T_zx \geq 0$ (the sign accord algorithm). This approach does not require computing the inverse matrix, but still the original Murty's form is to be recommended in the general case. Clearly, the sign accord algorithm cannot take more than 2^n steps; following Murty's example [14], we shall show here that this upper estimate can be achieved. For each $n \geq 1$ define an interval matrix A_n^I by

$$(A_n^I)_{ij} = \begin{cases} 1 \text{ if } i = j \\ [-2,2] \text{ if } j = i+1 \text{ and } 1 \leq i \leq n-1 \\ 0 \text{ otherwise.} \end{cases}$$

Now consider the system

$$A_n^I x = [f,f] . \qquad (3.1)$$

Theorem 3.1. For each $n \geq 1$, the sign accord algorithm takes exactly 2^n steps to compute x_e for the system (3.1) when started from $z = e$.

Proof. For the solution of the system $A_{ez} x = f$ we have

$$x_j = -1 - \sum_{i=1}^{n-j} 2^i z_{j+1} \cdots z_{j+i} \qquad (1 \leq j \leq n-1)$$

$$x_n = -1.$$

Hence, if the sign of z_k is reversed $(k \geq 2)$, then $x_k, \ldots, x_n$ remain unchanged and $x_1, \ldots, x_{k-1}$ take on values of opposite signs (since the last term $2^{n-j} z_{j+1} \cdots z_n$ prevails). Thus the sequence of vectors sgn $(T_z x)$ looks like this:

$-1,-1,-1,\ldots,-1$

$1,-1,-1,\ldots,-1$

$-1, 1,-1,\ldots,-1$

$1, 1,-1,\ldots,-1$

$-1,-1, 1,\ldots,-1$

...

...

If we write 0´s instead of -1´s, then, when read from right to left, these sequences will be just the binary numbers $0,1,\ldots,2^{n-1}$.
Thus the algorithm takes exactly 2^n steps before stopping with sgn $(T_z x) = e$. ∎

In [23] we showed that the sign accord algorithm can be started from an arbitrary $z \in Y$ without affecting its convergence. It is a good reason to believe that $z = \operatorname{sgn} d_y$ $(d_y = A_c^{-1} b_y)$ is a good choice. In fact, in ([21],p. 27) it was proved that if $C|d_y| < |d_y|$ (i.e. if A^I is "narrow"), then $\operatorname{sgn} x_y = \operatorname{sgn} d_y$, so that in this case the sign accord algorithm takes only one step. Nevertheless, in our above example we have $z = \operatorname{sgn} d_e = f$, $\operatorname{sgn}(T_z x) = ((-1)^{n-1}, \dots, 1, -1, 1)^T$, so that the algorithm takes more than 2^{n-2} steps.

In our paper [21] we described an iterative method for computing x_y. This method can be recommended for large-size examples with small values of $\rho(D)$ $(D = |A_c^{-1}|\Delta)$; small-size examples can be usually solved more quickly by finite algorithms. However, for the above example (where $\rho(D) = 0$) we have $x_e^0 = f$, $x_e^1 = x_e$.

4. A linear programming method for computing x_y

In this section we are going to show that the equation

$$A_{ye}x^+ - A_{yf}x^- = b_y$$

can be solved as a linear programming problem

$$\min \{ c^T(x_1 - x_2);\ A_{ye}x_1 - A_{yf}x_2 = b_y,\ x_1 \geqslant 0,\ x_2 \geqslant 0 \} \qquad (4.1)$$

if c is chosen so that the condition

$$c^T A^{-1} T_y > 0^T \quad \text{for each } A \in A^I \qquad (4.2)$$

be satisfied. This approach is similar to that of Mangasarian [11], but the proof is different.

Theorem 4.1. Let A^I be regular and let (4.2) hold. Then the problem (4.1) has a finite optimum and for each optimal solution x_1^*, x_2^* of (4.1) we have $x_y = x_1^* - x_2^*$.

Proof. Consider the primal problem

$$\min \left\{ c^T(x_1 - x_2) \,;\, \begin{array}{ll} (\underline{A}x_1 - \overline{A}x_2)_i \geqslant \overline{b}_i & (y_i = 1) \\ (-\overline{A}x_1 + \underline{A}x_2)_i \geqslant -\underline{b}_i & (y_i = -1) \\ x_1 \geqslant 0,\ x_2 \geqslant 0 \end{array} \right\} .$$

It can be easily verified that the dual problem reads

$$\max \left\{ b_y^T T_y p;\ A_{ye}^T T_y p \leqslant c,\ A_{yf}^T T_y p \geqslant c,\ p \geqslant 0 \right\} .$$

If p satisfies the dual constraints, then $|A_c^T T_y p - c| \leqslant \Delta^T |T_y p|$, hence theorem 1.1 gives that $A^T T_y p = c$ for some $A \in A^I$, so that $p^T = c^T A^{-1} T_y > 0^T$ due to (4.2). Thus both the primal and dual problem are feasible, hence the duality theorem assures the existence of an optimal solution x_1^*, x_2^* to (4.1) , which, since the dual optimal solution p^* is positive, must satisfy $A_{ye}x_1^* - A_{yf}x_2^* = b_y$. Assume that $x_{1j}^* > 0$, $x_{2j}^* > 0$ for some j with $\Delta_{.j} \neq 0$ (the j-th column of Δ) . Then the complementary slackness conditions give $(A_{ye}^T T_y p^*)_j = (A_{yf}^T T_y p^*)_j = c_j$ implying $\Delta_{.j}^T p^* = 0$, a contradiction. Hence $x_{1j}x_{2j} = 0$ for each j with $\Delta_{.j} \neq 0$, thus $A_{ye}(x_1^* - x_2^*)^+ - A_{yf}(x_1^* - x_2^*)^- = b_y$, which gives $x_y = x_1^* - x_2^*$. ■

Baumann [2] arrived at the same condition (4.2) in another context and described a 3x3 example such that for $y = e$, (4.2) does not hold for any $c \in R^n$. Thus the method described is seemingly not general, but shows an interesting connection with linear programming.

5. An estimating algorithm

If n is large, then the computation of the exact interval solution $x^I = [\underline{x},\overline{x}]$ may be too costly; in such cases, an estimate of x^I could suffice (cf. Neumaier [15]).

An algorithm for simultaneous estimating x^I and $(A^I)^{-1}$ can be derived from lemma 1 in [22] . The algorithm, starting with $A_c x = b^I$, goes along all the pairs $(i,j) \in N \times N$ and at each step records the effect of A_{ij} being "made free". In its most elaborate form, paying attention to the estimates to grow as slowly as possible, it seems to give good results.

We leave it on an interested reader to derive and test that algorithm himself (herself) .

6. Regularity and P-matrices

In [23] , we proved this "theorem A": If A^I is regular, then for each $A_1, A_2 \in A^I$, both $A_1^{-1}A_2$ and $A_1 A_2^{-1}$ are P-matrices. In this section, we give some consequences of this result.

<u>Theorem 6.1.</u> Let A^I be regular. Then for each $A_1, A_2 \in A^I$ we have:

(i) there exist $x_1 > 0$, $x_2 > 0$ such that $A_1 x_1 = A_2 x_2$,

(ii) there exist $x_1 > 0$, $x_2 > 0$ such that $A_1^{-1} x_1 = A_2^{-1} x_2$,

(iii) if $A_1 x_1 = A_2 x_2$ for some $x_1 \neq 0$, $x_2 \neq 0$, then $x_{1i} x_{2i} > 0$ for some $i \in N$,

(iv) if $A_1^{-1} x_1 = A_2^{-1} x_2$ for some $x_1 \neq 0$, $x_2 \neq 0$, then $x_{1i} x_{2i} > 0$ for some $i \in N$.

Proof. The proof follows from the above-quoted theorem A and from the results by Gale and Nikaido [8] : if A is a P-matrix, then (a) there is an $x > 0$ with $Ax > 0$, (b) if $x \neq 0$, then $x_i(Ax)_i > 0$ for some i. ■

For any $z \in Y$, let us call the orthants $\{x;\ T_z x \geq 0\}$, $\{x;\ T_z x \leq 0\}$ opposite. Notice that b is fixed in the next assertion.

Corollary 6.1. Let A^I be regular and let $b \neq 0$. Then the solution set of the system $A^I x = b$ cannot intersect two opposite orthants simultaneously.

Proof. Follows from the assertion (iii) of theorem 6.1. ■

Theorem 6.2. Let A^I be regular, $A_1, A_2 \in A^I$. Then each real eigenvalue of $A_1^{-1}(A_1 - A_2)$ is less than 1.

Proof. Assume there is a real eigenvalue t with $t \geq 1$. Then from $A_1^{-1}(A_1 - A_2)x = tx$, $x \neq 0$, we have $(A_1 + t^{-1}(A_2 - A_1))x = 0$, hence the matrix $A_1 + t^{-1}(A_2 - A_1) \in A^I$ is singular, a contradiction. ■

Let us denote by $\rho_0(A)$ the maximum of real eigenvalue moduli of A. If no real eigenvalue exists, we put $\rho_0(A) = 0$.

Corollary 6.2. Let A^I be regular. Then $\rho_0(A_c^{-1}\Delta_0) < 1$ for each Δ_0 with $|\Delta_0| \leq \Delta$.

Proof. Follows from theorem 6.2 applied to (a) $A_1 = A_c$, $A_2 = A_c - \Delta_0$, (b) $A_1 = A_c$, $A_2 = A_c + \Delta_0$. ■

Next we prove a regularity criterion (seemingly not of practical value). Let $D_{yz} = A_c^{-1} T_y \Delta T_z$ for $y, z \in Y$.

Theorem 6.3. A^I is regular if and only if $\rho_0(D_{yz}) < 1$ for each $y, z \in Y$.

Proof. The "only if" part follows from corollary 6.2. The proof of the "if" part is the same as in [3] . ■

Theorem 6.4. Let A^I be regular. Then for each Δ_o with $|\Delta_o| \leq \Delta$ we have $|A_c^{-1}\Delta_o|_{ii} < 1$ for each $i \in N$.

Proof. Put $A = A_c - \Delta_o \in A^I$. Since all diagonal entries of the P-matrix $A_c^{-1}A$ are positive and $A_c^{-1}\Delta_o = E - A_c^{-1}A$, we have $(A_c^{-1}\Delta_o)_{ii} < 1$ for each $i \in N$. To complete the proof, it suffices to apply the result just obtained to $-\Delta_o$. ■

Next we give two sufficient singularity conditions. Let $D_o = A_c^{-1}\Delta$.

Corollary 6.3. Let $|D_o|_{ii} \geq 1$ for some $i \in N$. Then A^I is singular.

Proof. Follows from theorem 6.4 for $\Delta_o = \Delta$. ■

Corollary 6.4. Let some of the matrices $\underline{A}\overline{A}^{-1}$, $\overline{A}\underline{A}^{-1}$, $\underline{A}^{-1}\overline{A}$, $\overline{A}^{-1}\underline{A}$ have a nonpositive diagonal element. Then A^I is singular.

Proof. If A^I were regular, then each of the four matrices would have positive diagonal in view of theorem A.■

In [23] , we proved this characterization $(A_y = A_{ye}^{-1}A_{yf})$:

Theorem 6.5. A^I is regular if and only if A_y is a P-matrix for each $y \in Y$.

The "only if" part follows from theorem A. To prove the "if" part, we need not use theorem 0 of [23] , but instead we may apply the following theorem 6.6; notice that theorems 6 and 8 in [22] can be also proved as its direct consequences:

Theorem 6.6. The following assertions are mutually equivalent:

(i) A^I is regular,

(ii) $A_{ye}x^+ - A_{yf}x^- = e_j$ has a solution for each $y \in Y, j \in N$,

(iii) $A_{ye}x_1 - A_{yf}x_2 = e_j$ has a nonnegative solution for each $y \in Y$, $j \in N$.

Proof. (i) $\Rightarrow$ (ii): from theorem 1.3. (ii) $\Rightarrow$ (iii): obvious. (iii) $\Rightarrow$ (i): assume that A^I is singular, then $A^T p = 0$ for some $A \in A^I$, $p \neq 0$; assume w.l.g. that $p_j < 0$ for some $j \in N$. Let $y = -\mathrm{sgn}\, p$, then $A_{ye}^T p \geq 0$, $A_{yf}^T p \leq 0$, $p_j < 0$, hence the system $A_{ye}x_1 - A_{yf}x_2 = e_j$ cannot have a nonnegative solution due to Farkas lemma. ■

Corollary 6.5. Let A_y be positive definite for each $y \in Y$. Then A^I is regular.

Proof. Follows from theorem 6.5 since each positive definite matrix is a P-matrix [8] . ■

The converse implication is, however, not true:

Example 6.1. The interval matrix

$$A^I = \begin{pmatrix} [-1,3] & 1 \\ 1 & 0 \end{pmatrix}$$

is obviously regular, but none of the matrices A_y $(y \in Y)$ is positive definite.

Finally we show that theorem 6.5 will not remain true if we replace the matrices $A_y = A_{ye}^{-1} A_{yf}$ by matrices $A_{ye} A_{yf}^{-1}$:

Example 6.2. The interval matrix

$$A^I = \begin{pmatrix} [1.5,3.5] & [-0.5,1.5] \\ [0.5,2.5] & [1.5,3.5] \end{pmatrix}$$

is singular but all the matrices $A_{ye} A_{yf}^{-1}$ $(y \in Y)$ are P-matrices.

7. Interval inverse representation

In [23] we proved that if A^I is regular, then for each $y \in Y$, $j \in N$ there exists a $z \in Y$ such that $T_z(A_{yz}^{-1})_{.j} \geqslant 0$. Generally this z need not be unique, but $(A_{yz}^{-1})_{.j}$ does, since obviously $(A_{yz}^{-1})_{.j}$ is equal to x_y for the system $A^I x = e_j$. Hence if we define $z_{yj} = \operatorname{sgn}(A_{yz}^{-1})_{.j}$, then z_{yj} is correctly defined and satisfies $T_{z_{yj}}(A_{yz_{yj}}^{-1})_{.j} \geqslant 0$. Now let $Q = \{(y, z_{yj});\ y \in Y,\ j \in N\}$. Then we have the following theorem, which is similar to theorem 6 in [22], but works with A_{yz}^{-1}'s instead of with obscure matrices B_y:

<u>Theorem 7.1.</u> Let A^I be regular. Then for each $A \in A^I$ there exist nonnegative diagonal matrices L_{yz} $((y,z) \in Q)$ satisfying $\sum_Q L_{yz} = E$ such that

$$A^{-1} = \sum_Q A_{yz}^{-1} L_{yz} \tag{7.1}$$

holds.

<u>Proof.</u> Fix a $j \in N$. Since $(A_{yz_{yj}}^{-1})_{.j}$ are just the vectors x_y for the system $A^I x = e_j$, from theorem 1.3 we have $(A^{-1})_{.j} = \sum_Y \lambda_{yj} (A_{yz_{yj}}^{-1})_{.j}$ for some nonnegative λ_{yj}, $\sum_Y \lambda_{yj} = 1$. Now we obtain the desired result when defining the diagonal matrices L_{yz} by $(L_{yz_{yj}})_{jj} = \lambda_{yj}$ $(y \in Y, j \in N)$. ∎

<u>Corollary 7.1.</u> Let A^I be regular. Then for the exact interval inverse $(A^I)^{-1} = [\underline{B}, \overline{B}]$ we have

$$\begin{aligned} \underline{B} &= \min_Q A_{yz}^{-1} \\ \overline{B} &= \max_Q A_{yz}^{-1} . \end{aligned} \tag{7.2}$$

Proof. Immediate, since for each $A \in A^I$ and each $i,j \in N$, $(A^{-1})_{ij}$ is a convex combination of $(A_{yz}^{-1})_{ij}$, $(y,z) \in Q$. ■

If A^I is inverse-stable, then $Q = \{(y,z_j);\ y \in Y,\ z_j = \mathrm{sgn}(A_c^{-1})_{.j},\ j \in N\}$. If A^I is positively invertible, then $Q = \{(y,e);\ y \in Y\}$. In the general case the structure of Q is difficult to judge; but notice that (7.1), (7.2) remain true if Q is replaced by YxY. Also, not each matrix of the form $\sum_Q A_{yz}^{-1} L_{yz}$ is equal to some A^{-1}, $A \in A^I$, in the general case, for otherwise the solution set of each system $A^I x = e_j$ would be convex.

8. Nonnegatively invertible matrices

We shall give here an elementary proof of three necessary and sufficient conditions for nonnegative invertibility of an interval matrix. (iii) and (iv) are known from [10], [22] .

Theorem 8.1. The following assertions are mutually equivalent:

(i) A^I is nonnegatively invertible,

(ii) $\overline{A}^{-1} \geq 0$ and $\rho(\overline{A}^{-1}(\overline{A} - \underline{A})) < 1$,

(iii) $\underline{A}^{-1} \geq 0$ and $\overline{A}^{-1} \geq 0$,

(iv) A^I is regular and $\overline{A}^{-1} \geq 0$.

Proof. (i) $\Rightarrow$ (iii), (i) $\Rightarrow$ (iv): obvious. Denote $\tilde{D} = \overline{A}^{-1}(\overline{A} - \underline{A})$. (iv) $\Rightarrow$ (ii): it follows from theorem 6.2 that $\rho(\tilde{D}) = \rho_o(\tilde{D}) < 1$. (iii) $\Rightarrow$ (ii): we have $(E - \tilde{D})^{-1} = \underline{A}^{-1}\overline{A} = E + \underline{A}^{-1}(\overline{A} - \underline{A}) \geq 0$, hence $\rho(\tilde{D}) < 1$. (ii) $\Rightarrow$ (i): for each $A \in A^I$ we have $A = \overline{A}(E - \overline{A}^{-1}(\overline{A} - A))$; since $\rho(\overline{A}^{-1}(\overline{A} - A)) \leq \rho(\tilde{D}) < 1$, it follows that $A^{-1} = \left(\sum_0^\infty (\overline{A}^{-1}(\overline{A} - A))^j\right)\overline{A}^{-1} \geq 0$. ■

In proving (ii) $\Rightarrow$ (i), we achieved this result.

Corollary 8.1. Let A^I be nonnegatively invertible. Then for each $A \in A^I$ we have

$$A^{-1} = \left(\sum_{j=0}^{\infty} (\bar{A}^{-1}(\bar{A} - A))^j\right)\bar{A}^{-1} .$$

We may use this corollary to obtain a generalization of the result by Barth, Nuding [1] and Beeck [5] ($\underline{b} \geq 0$ is replaced by $\bar{A}^{-1}\underline{b} \geq 0$):

Theorem 8.2. Let A^I be nonnegatively invertible and let $\bar{A}^{-1}\underline{b} \geq 0$. Then for the exact interval solution $[\underline{x},\bar{x}]$ of $A^I x = b^I$ we have $\underline{x} = \bar{A}^{-1}\underline{b}$, $\bar{x} = \underline{A}^{-1}\bar{b}$.

Proof. For each $x = A^{-1}b$, $A \in A^I$, $b \in b^I$, corollary 8.1 gives $x \geq \bar{A}^{-1}\underline{b} \geq 0$. Thus $Ax = b$ implies $\underline{A}x \leq \bar{b}$ and $x \leq \underline{A}^{-1}\bar{b}$. Hence $\underline{x} = \bar{A}^{-1}\underline{b}$, $\bar{x} = \underline{A}^{-1}\bar{b}$. ■

In [22] it was shown that $\underline{x} = x_f$, $\bar{x} = x_e$ for positively invertible matrices. This result can be extended to inverse-column-stable matrices, i.e. to interval matrices satisfying $A^{-1}T_z > 0$ for some (fixed) $z \in Y$ and each $A \in A^I$. In this case we have $\underline{x} = x_{-z}$, $\bar{x} = x_z$. Obviously, A^I is inverse-column-stable iff the interval matrix $[T_z A_{ze}, T_z A_{zf}]$ is positively invertible, i.e. e.g. iff $(T_z A_{ze})^{-1} > 0$ and $(T_z A_{zf})^{-1} > 0$.

9. Nonnegative solutions

Nonnegative solutions play important role in certain applications, e.g. in interval linear programming. Obviously, the set of nonnegative solutions X_+ is described by $X_+ = \{x;\ \underline{A}x \leq \overline{b},\ \overline{A}x \geq \underline{b},\ x \geq 0\}$ (corollary 1.2). We shall focus our attention here on two problems: when (a) $X_+ = X$, (b) $X_+ \neq \emptyset$.

<u>Theorem 9.1.</u> Let A^I be regular. Then each solution $x \in X$ is nonnegative (positive) if and only if $A_{ye}^{-1}b_y \geq 0$ (> 0) for each $y \in Y$.

<u>Proof.</u> "Only if": obvious. "If": denote $\tilde{x}_y = A_{ye}^{-1}b_y$, then $A_{ye}\tilde{x}_y = b_y$ and $T_e\tilde{x}_y \geq 0$, hence $\tilde{x}_y = x_y$. Now the assertion follows from theorem 1.3. ■

Thus verifying nonnegativity of all solutions requires computation of 2^n vectors x_y; this, however, does not mean that one must solve 2^n systems $A_{ye}x = b_y$. In fact, if y and y' differ in just one entry, then $x_{y'}$ can be computed from x_y using lemma 1 in [22]. Thus we are left with the problem whether the set Y can be ordered in such a way that every two neighbouring vectors differ in just one entry. Such an ordering can be easily constructed by induction on n: (i) $1, -1$ is the ordering for $n = 1$, (ii) if $y_1, \ldots, y_{2^n}$ is the ordering for n, then $(y_1, 1), \ldots, (y_{2^n}, 1), (y_{2^n}, -1), \ldots, (y_1, -1)$ is the ordering for n+1.

Next we turn to the problem of existence of nonnegative solutions. The following theorem was proved in [18]. Regularity is not assumed; notice the difference in quantifiers:

<u>Theorem 9.2.</u> A system $A^Ix = b^I$ has a nonnegative solution if and only if for each p satisfying $A^Tp \geq 0$ for each $A \in A^I$ there exists a $b \in b^I$ such that $b^Tp \geq 0$.

This is a Farkas-type theorem. It has this consequence:

Corollary 9.1. A system $A^I x = b^I$ does not possess a nonnegative solution if and only if there exists a fixed linear combination of rows which, as applied to any system $Ax = b$, $A \in A^I$, $b \in b^I$, always produces an equation that does not possess a nonnegative solution.

Proof. If $A^I x = b^I$ does not possess a nonegative solution, then theorem 9.2 assures the existence of a vector p such that $A^T p \geqslant 0$, $b^T p < 0$ for any $A \in A^I$, $b \in b^I$. Taking a linear combination of rows with coefficients p_i $(i \in N)$, a system $Ax = b$ is brought to the equation $p^T Ax = b^T p$ that cannot have a nonnegative solution. Conversely, if such a p exists, then the assumption of existence of a nonnegative solution to some $Ax = b$ leads to the existence of a nonnegative solution of the equation $p^T Ax = b^T p$, a contradiction.■

10. Interval linear programming

An interval linear programming problem is a problem of the form

$$\max \{c^T x;\ Ax = b,\ x \geqslant 0\} \tag{10.1}$$

where A,b,c are not exactly known: $A \in \tilde{A}^I$, $b \in b^I$, $c \in \tilde{c}^I$ ($\tilde{A}^I$ is of size nxm; the symbol A^I is reserved for certain square submatrix of $\tilde{A}^I$) . We shall assume that for each such A,b,c the problem (10.1) has a finite optimum (necessary and sufficient conditions were given in [17]); let us denote the optimal value of (10.1) by f(A,b,c) . Several problems arise in connection with this formulation:

(a) Compute the upper bound of optimal values

$$\bar{h} = \max\{f(A,b,c);\ A \in \tilde{A}^I,\ b \in b^I,\ c \in \tilde{c}^I\} .$$

It can be easily seen that $\bar{h}$ can be computed by a linear programming method:

$$\bar{h} = \max\{\bar{\tilde{c}}^T x\ ;\ \underline{\tilde{A}}x \leq \bar{b},\ \bar{\tilde{A}}x \geq \underline{b},\ x \geq 0\}. \qquad (10.2)$$

In [18] , a duality theory for this problem was developed (notice the additional duality between "weak" and "strong" solutions of (P) and (D) there; it can be also extended to inequalities provided nonnegativity of the primal solution is retained) resulting in the optimality criterion 3.4 (in [18]) which implies this algorithm, working, in contrast to (10.2) , with problem of original size:

<u>Algorithm 10.1.</u>

0. Set $y = e$.
1. Compute $h = \max\{\bar{\tilde{c}}^T x;\ \tilde{A}_{ye} x = b_y,\ x \geq 0\}$ and a dual optimal solution p.
2. If $T_y p \geq 0$, terminate with $\bar{h} = h$.
3. Otherwise select a j with $y_j p_j < 0$, set $y_j = -y_j$ and go to step 1.

The following theorem was proved in [19] :

<u>Theorem 10.1.</u> Let $\delta > 0$ and let each nonnegative solution to $\tilde{A}^I x = b^I$ be nondegenerate (i.e. it has at least n positive entries) . Then algorithm 10.1 is finite.

Various improvements and schemes for performing this algorithm in a tableau can be given. One scheme (with fixed-size tableau) was suggested in [19] , another (with floating-size tableau) was recently proposed by Mráz [12] .

(b) Compute the lower bound of optimal values

$$\underline{h} = \min\left\{ f(A,b,c);\ A \in \tilde{A}^I,\ b \in b^I,\ c \in \tilde{c}^I \right\} .$$

Taking A,b,c fixed and letting p be a dual optimal solution to (10.1), we can easily obtain from the duality theorem that $f(A,b,c) \geq$ $\geq f(\tilde{A}_{ye}, b_y, \underline{\tilde{c}})$, where $y = -\operatorname{sgn} p$. Hence

$$\underline{h} = \min_{Y} f(\tilde{A}_{ye}, b_y, \underline{\tilde{c}}) ,$$

which would require solving 2^n problems (10.1). Due to the lack of an optimality criterion, no analogon of algorithm 10.1 and theorem 10.1 is known as yet, as far as we know.

(c) Compute the upper and lower bounds of optimal solutions:

$$\underline{x}_i = \min\left\{ x_i;\ x \in X^* \right\}$$
$$\overline{x}_i = \max\left\{ x_i;\ x \in X^* \right\} \qquad (i \in N),$$

where X^* is the set of all optimal solutions of all the problems (10.1) for $A \in \tilde{A}^I$, $b \in b^I$, $c \in \tilde{c}^I$. In the general case this seems to be a difficult problem.

We shall show that the problems (a), (b), (c) can be rather efficiently solved in a special case of basis-stable problems studied earlier by Krawczyk [9] and Beeck [6] . An interval linear programming problem is called basis-stable with basis B if for each $A \in \tilde{A}^I$, $b \in b^I$, $c \in \tilde{c}^I$, the problem (10.1) has a unique non-degenerate optimal basic solution with basis B. We shall assume without loss of generality that B is formed by the first n columns, so that $\tilde{A}^I$, $\tilde{c}^I$ can be decomposed into basic and nonbasic parts: $\tilde{A}^I = (A^I, A_N^I)$, $(\tilde{c}^I)^T = (c^{IT}, c_N^{IT})$. Now the above problems (a), (b), (c) can be formulated as problems of the form

$$\max\left\{ c^T x;\ x \in X \right\} \qquad (10.3)$$
$$\min\left\{ c^T x;\ x \in X \right\}, \qquad (10.4)$$

where X is the solution set of $A^I x = b^I$ (A^I is square!). In fact, (a) is of the form (10.3) with $c = \bar{c}$, (b) of the form (10.4) with $c = \underline{c}$, (c) of the form (10.3) or (10.4) with $c = e_i$, $i \in N$ (if $i > n$, then $\underline{x}_i = \bar{x}_i = 0$). Furthermore, we need not solve the problem (10.3) or (10.4) by the simplex method, but instead by the following algorithm (making use of the special structure of our problem), which was developed by the author in collaboration with the undergraduate student I. Bureš [7] (usually, solving one problem (10.1) must precede to determine the basis):

Algorithm 10.2. (for solving (10.3)).

0. Set $y = e$.
1. Solve $A_{ye} x = b_y$, $A^T_{ye} p = c$.
2. If $T_y p \geq 0$, terminate. x is the optimal solution of (10.3).
3. Otherwise select a j with $y_j p_j < 0$, set $y_j = -y_j$ and go to step 1.

In the case of solving (10.4), the stopping rule would be $T_y p \leq 0$, hence in step 3 we would search for a j with $y_j p_j > 0$. Also the following theorem holds for (10.4) with this change:

Theorem 10.2. Let the original problem be basis-stable and let the basic system $A^I x = b^I$ be normal. Then the algorithm terminates in a finite number of steps with an optimal solution to (10.3).

Proof. Notice that the assumptions of theorem 2.4 are met, so that the algorithm goes along the edges of the convex polyhedron X. If x_y, $x_{y'}$ are neighbouring extremal points with y, y′ differing only in the j-th entry, then from lemma 1 in [22] we obtain $x_{y'} = x_y - \beta y_j (A^{-1}_{ye})_{.j}$ for some $\beta > 0$, so that $c^T x_{y'} = c^T x_y - \beta y_j p_j$. Hence at each step the objective value increases, which assures finiteness. If $T_y p \geq 0$, then the objective value does not increase along any edge emanating from x_y, hence $c^T x_y$ is optimal due to the convexity of X. ∎

With the help of lemma 1 in [22], the algorithm can be performed in a tableau form, storing A_{ye}^{-1}, x, p and y at each step. Details are left to the reader, as well as deriving a verifiable sufficient condition for basis-stability.

11. Problems

Several problems I have failed to solve are listed here. Maybe some of them are easy, in which case I apologize to the reader, who, I hope, will be more successful in solving them than I did:

Problem 1. Describe Conv X by a system of linear inequalities in terms of input data A^I, b^I.

Problem 2. Does an analogue of theorem 2.4 hold for the set Conv X in the general case?

Problem 3. Develop a linear programming procedure for computing $\underline{x}_i$, $\overline{x}_i$ directly. Can corollary 1.4 be used for this purpose?

Problem 4. Develop an efficient nonlinear programming procedure for computing $\underline{x}_i$, $\overline{x}_i$ directly.

Problem 5. Is algorithm [20, p. 51] finite for an arbitrary regular matrix A^I?

Problem 6. Does there exist a necessary and sufficient condition for the iterative method [21, Eq.(11)] to converge? ($\rho(D) < 1$ is sufficient, but not necessary; what about $\max_{y,z \in Y} \rho(A_c^{-1} T_y \Delta T_z) < 1$?)

References

[1] W. Barth, E. Nuding, Optimale Lösung von Intervallgleichungssystemen, Computing 12(1974), 117-125

[2] M. Baumann, Solutions of Systems of Interval Linear Equations (in Czech) , thesis, Charles University, Prague 1983

[3] M. Baumann, A Regularity Criterion for Interval Matrices, in: "Herrn Professor Dr. Karl Nickel zum 60. Geburtstag gewidmet" (J. Garloff, A. Neumaier, D. Norbert, A. Schäfer, Eds.), Freiburg 1984

[4] H. Beeck, Charakterisierung der Lösungsmenge von Intervallgleichungssystemen, ZAMM 53(1973), T181-T182

[5] H. Beeck, Zur scharfen Aussenabschätzung der Lösungsmenge bei linearen Intervallgleichungssystemen, ZAMM 54(1974), T208-T209

[6] H. Beeck, Linear Programming with Inexact Data, research report TUM-ISU-7830, Munich 1978

[7] I. Bureš, Linear Programming Problems with Interval Data (in Czech) , thesis, Charles University, Prague 1983

[8] D. Gale, H. Nikaido, The Jacobian Matrix and Global Univalence of Mappings, Math. Annalen 159(1965), 81-93

[9] R. Krawczyk, Fehlerabschätzung bei linearer Optimierung, in: Interval Mathematics (K. Nickel, Ed.) , Springer-Verlag, Berlin-Heidelberg-New York 1975, 215-222

[10] J. Kuttler, A Fourth-Order Finite-Difference Approximation for the Fixed Membrane Eigenproblem, Math. of Computation 25(1971), 237-256

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

[12] F. Mráz, An Algorithm for Solving Interval Linear Programming Problems (in Czech), manuscript, České Budějovice 1983

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

[14] K. Murty, Computational Complexity of Complementary Pivot Methods, Mathematical Programming Study 7(1978), 61-73

[15] A. Neumaier, New Techniques for the Analysis of Linear Interval Equations, Freiburger Intervall-Berichte 83/4

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

[17] J. Rohn, Strong Solvability of Interval Linear Programming Problems, Computing 26(1981), 79-82

[18] J. Rohn, Duality in Interval Linear Programming, in: Interval Mathematics 1980 (K.L.E. Nickel, Ed.), Academic Press, New York 1980, 521-529

[19] J. Rohn, Duality in Interval Linear Programming, an extended manuscript of [18], partially unpublished; presented at the International Symposium on Interval Mathematics, Freiburg, May 1980

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

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

[22] J. Rohn, Solving Interval Linear Systems, in: "Herrn Professor Dr. Karl Nickel zum 60. Geburtstag gewidmet" (J. Garloff, A. Neumaier, D. Norbert, A. Schäfer, Eds.), Freiburg 1984; reprinted in Freiburger Intervall-Berichte

[23] J. Rohn, Proofs to "Solving Interval Linear Systems", Freiburger Intervall-Berichte

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