BIT 30 (1989), 161–165

ON NONCONVEXITY OF THE SOLUTION SET OF A SYSTEM OF LINEAR INTERVAL EQUATIONS

J. ROHN

Faculty of Mathematics and Physics, Charles University, Malostranské nám. 25, 11800 Prague, Czechoslovakia

Abstract.

We give necessary and sufficient conditions for the solution set of a system of linear interval equations to be nonconvex and derive some consequences.

AMS Subject Classifications: 15A06, 52A99, 65G10.

Key words: linear interval equations, solution set, nonconvexity.

1. Introduction.

Since the early 1960's the problem of uncertainties in the data in systems of linear equations has been studied with the help of linear interval equations. These linear interval equations arise in a natural manner when it is assumed that the coefficients and the right hand sides of a linear system $A_0x = b_0$ are enclosed in intervals which are assumed to be mutually independent. Instead of the original system one is therefore led to consider the system of linear interval equations

$$A^I x = b^I \tag{1}$$

where $A^I = \{A; A_c - \Delta \leqslant A \leqslant A_c + \Delta\}$ is an $n \times n$ interval matrix and $b^I = \{b; b_c - \delta \leqslant b \leqslant b_c + \delta\}$ is an interval vector; here A_c is the center of A^I, b_c the center of b^I and the radius matrix Δ as well as the radius vector δ are nonnegative. The solution set X is defined as the solution range over the intervals, that is,

$$X = \{x; Ax = b, A \in A^I, b \in b^I\}$$

because of the assumption of the independence of the coefficients. It is easy to give examples of solution sets that are nonconvex. Even so it was shown by Oettli and Prager [7] that the solution set can be characterized in a surprisingly simple

Received February 1989. Revised April 1989.

manner. They showed that for arbitrary A^I, b^I the set X can be described by

$$X = \{x; |A_c x - b_c| \leqslant \Delta |x| + \delta\} \tag{2}$$

where $|x|$, the absolute value of x, is defined by $|x|_i = |x_i|$ $(i = 1, \ldots, n)$. The source of the nonconvexity of X is the term $|x|$ appearing on the right hand side of the inequality in (2). The nonconvexity disappears, however, when X is entirely contained in an orthant since it was observed by Oettli [6] that the intersection of X with each orthant is a convex polytope. Some nice examples illustrating the (sometimes complicated) structure of the solution set X are given e.g. in [1], [3], [4], [5], and further discussions and results on linear interval systems can be found in Alefeld and Herzberger [10].

In this paper we will give a necessary and sufficient condition for X to be nonconvex (Theorem 2 below) as well as some of the consequences of the condition. One of the consequences is that if X is convex, if the radius matrix Δ is positive and if a rather weak additional condition is satisfied, then X is already included in a single orthant (Corollary 2). This result is almost a converse of the assertion by Oettli quoted above.

2. Main result.

We begin with some notations. Denote $Y = \{y \in R^n; |y_j| = 1 \text{ for each } j\}$, so that Y consists of 2^n vectors. For each $y \in Y$, let $T_y = \operatorname{diag}\{y_1, \ldots, y_n\}$ denote the diagonal matrix with diagonal vector y. Our characterization of nonconvexity of X will be based on the following convex hull theorem proved in [9]; a square interval matrix A^I is called regular if each $A \in A^I$ is nonsingular.

THEOREM 1: *Let A^I be regular. Then for each $y \in Y$, the nonlinear equation*

$$A_c x - b_c = T_y(\Delta |x| + \delta) \tag{3}$$

has exactly one solution x_y, $x_y \in X$, and there holds

$$\operatorname{Conv} X = \operatorname{Conv}\{x_y; y \in Y\}. \tag{4}$$

A general finite method for computing x_y for a given $y \in Y$ was given in [9]. Here we shall use the vectors x_y to characterize nonconvexity:

THEOREM 2: *Let A^I be a regular $n \times n$ interval matrix. Then the solution set X of (1) is nonconvex if and only if there exist y, $z \in Y$ and $i, j \in \{1, \ldots, n\}$ such that $y_i = z_i$, $(x_y)_j(x_z)_j < 0$ and $\Delta_{ij} > 0$.*

PROOF: (a) We shall first prove the "if" part. Assuming that y, z, i and j with the properties listed exist, take real numbers $\lambda > 0$, $\mu > 0$ with $\lambda + \mu = 1$ and put $x = \lambda x_y + \mu x_z$. Then $|x|_j < \lambda |x_y|_j + \mu |x_z|_j$, while $|x|_k \leqslant \lambda |x_y|_k + \mu |x_z|_k$ for each

$k \neq j$. Since x_y and x_z satisfy

$$(A_c x_y - b_c)_i = y_i(\Delta |x_y| + \delta)_i$$
$$(A_c x_z - b_c)_i = z_i(\Delta |x_z| + \delta)_i = y_i(\Delta |x_z| + \delta)_i$$

due to (3), we obtain, using the positivity of Δ_{ij}, that

$$|A_c x - b_c|_i = (\Delta(\lambda |x_y| + \mu |x_z|) + \delta)_i > (\Delta |x| + \delta)_i,$$

which in view of (2) means that $x \notin X$. Since $x_y, x_z \in X$ and x belongs to the segment connecting them, we conclude that X is nonconvex.

(b) To prove the "only if" part of the theorem, assume on the contrary that for each $y, z \in Y$ and each $i, j \in \{1, \ldots, n\}$, $y_i = z_i$ and $\Delta_{ij} > 0$ imply $(x_y)_j(x_z)_j \geqslant 0$. We shall prove that in this case each convex combination of vectors x_y belongs to X. This, in the light of (4), will imply that $\operatorname{Conv} X \subset X$, proving that X is convex. So let $x = \sum_{y \in Y} \lambda_y x_y$, where $\lambda_y, y \in Y$, are nonnegative real numbers satisfying $\sum_{y \in Y} \lambda_y = 1$. Then from (3) we have

$$\begin{aligned}(A_c x - b_c)_i &= \sum_{y \in Y} \lambda_y (A_c x_y - b_c)_i = \sum_{y \in Y} \lambda_y y_i (\Delta |x_y| + \delta)_i \\ &= \sum_{j=1}^{n} \Delta_{ij} \left(\sum_{y \in Y} \lambda_y y_i |x_y|_j \right) + \sum_{y \in Y} \lambda_y y_i \delta_i\end{aligned}$$

and using our assumption that $y_i = z_i$ and $\Delta_{ij} > 0$ imply $(x_y)_j(x_z)_j \geqslant 0$, we obtain

$$(A_c x - b_c)_i = \sum_{j=1}^{n} \Delta_{ij} \left(\left| \sum_{\substack{y \in Y \\ y_i = 1}} \lambda_y x_y \right|_j - \left| \sum_{\substack{y \in Y \\ y_i = -1}} \lambda_y x_y \right|_j \right) + \sum_{y \in Y} \lambda_y y_i \delta_i.$$

Taking absolute values we have

$$|A_c x - b_c|_i \leqslant \sum_{j=1}^{n} \Delta_{ij} \left| \sum_{y \in Y} \lambda_y x_y \right|_j + \delta_i = (\Delta |x| + \delta)_i$$

for each $i \in \{1, \ldots, n\}$ and hence $|A_c x - b_c| \leqslant \Delta |x| + \delta$. This implies $x \in X$ in view of (2), and hence $\operatorname{Conv} X \subset X$. ∎

3. Consequences

First, from the proof of Theorem 2 we obtain:

COROLLARY 1: *Let A^I be regular. Then X is nonconvex if and only if there exist $y, z \in Y$ with $y \neq -z$ and $x_y \neq x_z$ such that no point of the segment connecting x_y with x_z, except the endpoints, belongs to X.*

PROOF: The "if" part is obvious. Conversely, if X is nonconvex, there exist y, z, i, j such that $y_i = z_i$, $(x_y)_j(x_z)_j < 0$ and $\Delta_{ij} > 0$ and according to part (a) of the proof of Th. 2, no point of the form $\lambda x_y + \mu x_z$, $\lambda > 0, \mu > 0, \lambda + \mu = 1$, belongs to X. ■

Next theorem explains the reason for the condition $y \neq -z$ in Corollary 1:

THEOREM 3: *Let A^I be regular. Then for each $y \in Y$, the whole segment connecting x_y with x_{-y} belongs to X.*

PROOF: Let $y \in Y$ and let $x = \lambda x_y + \mu x_{-y}$ for some $\lambda \geqslant 0, \mu \geqslant 0, \lambda + \mu = 1$. Then from (3) we get $|A_c x - b_c| = |T_y(\Delta(|\lambda x_y| - |\mu x_{-y}|) + (\lambda - \mu)\delta)| \leqslant \Delta|x| + \delta$; hence $x \in X$ according to (2). ■

We shall now turn to a special case of interval matrices

$$A^I = \{A; A_c - \Delta \leqslant A \leqslant A_c + \Delta\}$$

satisfying $\Delta > 0$ (componentwise); they are sometimes called thick interval matrices. Here the situation becomes more clear:

THEOREM 4: *Let A^I be a regular interval matrix satisfying $\Delta > 0$. Then X is nonconvex if and only if there exist $y, z \in Y$, $y \neq -z$, such that $(x_y)_j(x_z)_j < 0$ for some $j \in \{1, \ldots, n\}$.*

PROOF: The "only if" assertion follows directly from Theorem 2. Conversely, if $(x_y)_j(x_z)_j < 0$ for some $y, z \in Y$, $y \neq -z$, then there exists an i with $y_i = z_i$ and since $\Delta_{ij} > 0$, Theorem 2 applies. ■

In the introductory section we quoted Oettli's result stating that if X is part of one orthant, then X is convex. We shall now show that under additional assumptions this result can also be reversed:

COROLLARY 2: *Let A^I be regular with $\Delta > 0$, let X be convex and let there exist a $y \in Y$ such that both x_y and x_{-y} belong to the same orthant and at least one of them has all entries nonzero. Then X is part of a single orthant.*

PROOF: Assume that all entries of x_y are nonzero. Since X is convex, it follows from Theorem 4 that for each $z \in Y$, $z \neq -y$, there holds $(x_y)_j(x_z)_j \geqslant 0$ for each j, so that each such x_z belongs to the same orthant as x_y. Since this is also true for the remaining vector x_{-y} due to the assumption, we obtain from Theorem 1, Eq. (4) that the whole solution set X is also part of that very orthant. ■.

In fact, without our additional assumptions, even in the case $\Delta > 0$ the solution set X can be convex and still intersect all the orthants, as the following example shows:

EXAMPLE: Let $A_c = \begin{pmatrix} 1.5 & 1.5 \\ 1.5 & -1.5 \end{pmatrix}, \quad \Delta = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}$

and $b_c = (0,0)^T$, $\delta = (1,1)^T$. Then the solution set X of (1) is the square with vertices $(1,0)^T, (0,1)^T, (-1,0)^T, (0,-1)^T$.

In the last theorem we give another property of the solution set X, valid, however, only in the case $\delta = 0$ (i.e. when the right hand side in (1) reduces to one vector). For any $z \in Y$, let us call the orthants $\{x; z_j x_j \geqslant 0 \text{ for each } j\}$ and $\{x; z_j x_j \leqslant 0 \text{ for each } j\}$ opposite.

THEOREM 5: *Let A^I be regular and let $b \neq 0$. Then the solution set of the interval linear system $A^I x = b$ cannot intersect simultaneously two opposite orthants.*

PROOF: Let $x_1, x_2 \in X$, so that $A_1 x_1 = b$, $A_2 x_2 = b$ for some $A_1, A_2 \in A^I$. Then $x_1 = A_1^{-1} A_2 x_2$ and $x_2 \neq 0$ due to $b \neq 0$. Since A^I is regular, $A_1^{-1} A_2$ is a P-matrix, as proved in [9], Theorem 1.2, and from the well-known characterization of P-matrices by Fiedler and Pták [2] we obtain that there exists a $j \in \{1, \ldots, n\}$ such that $(x_1)_j (x_2)_j > 0$. Hence x_1 and x_2 cannot belong to two opposite orthants. ■

A preliminary version of this paper appeared in [8].

Acknowledgement.

The author thanks a referee for suggestions that helped to improve the text of the introduction.

REFERENCES

[1] W. Barth and E. Nuding, *Optimale Lösung von Intervallgleichungssystemen*, Computing 12 (1974), 117–125.
[2] M. Fiedler and V. Pták, *On matrices with non-positive off-diagonal elements and positive principal minors*, Czech. Math. Journal 12 (1962), 382–400.
[3] D. Gay, *Solving interval linear equations*, SIAM J. Numer. Anal. 19 (1982), 858–870.
[4] E. Hansen, *On linear algebraic equations with interval coefficients*, in: *Topics in Interval Analysis* (E. Hansen, Ed.), Clarendon Press, Oxford 1969.
[5] K. Nickel, *Die Überschätzung des Wertebereichs einer Funktion in der Intervallrechnung mit Anwendungen auf lineare Gleichungssysteme*, Computing 18 (1977), 15–36.
[6] W. Oettli, *On the solution set of a linear system with inaccurate coefficients*, SIAM J. Numer. Anal. 2 (1965), 115–118.
[7] W. Oettli and 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.
[8] J. Rohn, *Interval linear systems*, Freiburger Intervall-Berichte 84/7, 33–58.
[9] J. Rohn, *Systems of linear interval equations*, to appear in Lin. Alg. Appls.
[10] G. Alefeld and J. Herzberger, *Introduction to Interval Computations*, Academic Press, New York 1983.