SIAM J. MATRIX ANAL. APPL.
Vol. 18, No. 3, pp. 693–705, July 1997

ON THE SHAPE OF THE SYMMETRIC, PERSYMMETRIC, AND SKEW-SYMMETRIC SOLUTION SET*

GÖTZ ALEFELD†, VLADIK KREINOVICH‡, AND GÜNTER MAYER§

Dedicated to Prof. Dr. Gerhard Maeß, Rostock, on the occasion of his 60th birthday.

Abstract. We present a characterization of the solution set S, the symmetric solution set S_{sym}, the persymmetric solution set S_{per}, and the skew-symmetric solution set S_{skew} of real linear systems $Ax = b$ with the $n \times n$ coefficient matrix A varying between a lower bound $\underline{A}$ and an upper bound $\overline{A}$, and with b similarly varying between $\underline{b}$, $\overline{b}$. We show that in each orthant the sets S_{sym}, S_{per}, and S_{skew} are, respectively, the intersection of S with sets, the boundaries of which are quadrics.

Key words. linear systems with perturbed input data, solution set of linear systems of equations, symmetric matrices, persymmetric matrices, skew-symmetric matrices, Oettli–Prager theorem, Fourier–Motzkin elimination, interval analysis

AMS subject classifications. 65G10, 65F05

PII. S0895479896297069

1. Introduction. Let $[A]$ be an $n \times n$ matrix with compact intervals as entries, let $[b]$ be a vector with n interval components, and let E be the $n \times n$ permutation matrix with ones in the northeast–southwest diagonal and zeros elsewhere. The purpose of this paper is to characterize the solution sets

$$S := \{x \in \boldsymbol{R}^n \mid Ax = b,\ A \in [A],\ b \in [b]\}, \tag{1.1}$$

$$S_{sym} := \{x \in \boldsymbol{R}^n \mid Ax = b,\ A = A^T \in [A] = [A]^T,\ b \in [b]\}, \tag{1.2}$$

$$S_{per} := \{x \in \boldsymbol{R}^n \mid Ax = b,\ EA = (EA)^T \in E[A] = (E[A])^T,\ b \in [b]\}, \tag{1.3}$$

$$S_{skew} := \{x \in \boldsymbol{R}^n \mid Ax = b,\ A = -A^T \in [A] = ([a]_{ij}) = -[A]^T,\ [a]_{ii} = 0 \text{ for } i = 1, \ldots, n,\ b \in [b]\} \tag{1.4}$$

by means of inequalities which show that in each fixed orthant O the solution set S is the intersection of finitely many half spaces, while $S_{sym} \cap O$, $S_{per} \cap O$, and $S_{skew} \cap O$ are the intersection of $S \cap O$ with finitely many sets, the boundaries of which are conic sections in $\boldsymbol{R}^n$. The characterization of $S \cap O$ was already given in [4], [5], [7], [11], [12], and others while the characterization of $S_{sym} \cap O$ in the two-dimensional case was derived in [4]. The technique there could not be transferred onto the general case in an obvious way. It was changed in [2], [3]. We will use here a different technique known as Fourier–Motzkin elimination, which is described, e.g., in [14].

*Received by the editors January 8, 1996; accepted for publication (in revised form) by N. J. Higham July 18, 1996.
http://www.siam.org/journals/simax/18-3/29706.html

†Institut für Angewandte Mathematik, Universität Karlsruhe, D-76128 Karlsruhe, Germany (goetz.alefeld@mathematik.uni-karlsruhe.de).

‡Department of Computer Science, University of Texas at El Paso, El Paso, TX 79968 (vladik@cs.utep.edu).

§Fachbereich Mathematik, Universität Rostock, D-18051 Rostock, Germany (guenter.mayer@mathematik.uni-rostock.de).

Note that we require

$$(1.5)\quad \begin{cases} \text{no additional condition on } [A] \text{ in the case of } S, \\ [A]=[A]^T \text{ in the case of } S_{sym}, \\ E[A]=(E[A])^T \text{ in the case of } S_{per}, \\ [A]=([a]_{ij})=-[A]^T \text{ with } [a]_{ii}=0,\ i=1,\dots,n \text{ in the case of } S_{skew}. \end{cases}$$

The restrictions in (1.5) are not severe. If $[A] \neq [A]^T$ in the case of S_{sym}, e.g., and if $[B]$ denotes the largest interval matrix in $[A]$ such that $[B]=[B]^T$ holds, then the matrices in $[A]\backslash[B]$ do not influence S_{sym}. Therefore, instead of $[A]$ the matrix $[B]$ would play the crucial role in characterizing S_{sym}.

We emphasize that $[A]$ is allowed to contain singular real matrices. The restriction $[a]_{ii}=0$, $i=1,\dots,n$ in the case of S_{skew} stems from the fact that a skew-symmetric matrix $A=(a_{ij}) \in \boldsymbol{R}^{n\times n}$ is defined by $A=-A^T$ which implies $a_{ii}=0$ for $i=1,\dots,n$. We also recall that this matrix is singular if n is odd. This can be seen from $\det A=\det(-A^T)=\det(-A)=(-1)^n \det A$. The condition $EA=(EA)^T$ for S_{per} characterizes a persymmetric matrix which is defined to be symmetric with respect to the northeast–southwest diagonal; cf. [6], e.g.

The sets in (1.1)–(1.4) occur when dealing with linear systems of equations, the input data of which are afflicted with tolerances (cf. [1], [10], or [13], e.g.). This is the case when data $\check{A}$, $\check{b}$ are perturbed by errors caused, e.g., by measurements or by a conversion from decimal to binary digits on a computer. Assume that these errors are known to be bounded by some quantities $\Delta A \in \boldsymbol{R}^{n\times n}$ and $\Delta b \in \boldsymbol{R}^n$ with nonnegative entries. Then it seems reasonable to accept a vector $\tilde{x}$ as the "correct" solution of $\check{A}x=\check{b}$ if it is in fact the solution of a perturbed system $\tilde{A}x=\tilde{b}$ with

$$\tilde{A} \in [A] := [\check{A}-\Delta A, \check{A}+\Delta A], \quad \tilde{b} \in [b] := [\check{b}-\Delta b, \check{b}+\Delta b].$$

The characterization of all such $\tilde{x}$ led Oettli and Prager [11] to their famous equivalence

$$(1.6)\qquad x \in S \Longleftrightarrow |\,\check{b}-\check{A}x\,| \le \Delta A|x| + \Delta b,$$

where $|v| := (|v_i|) \in \boldsymbol{R}^n$ for $v=(v_i) \in \boldsymbol{R}^n$. It relates the midpoint residual to the tolerances and to $|x|$ and was reformulated in [7] similarly as in the subsequent Theorem 3.4. Often $\check{A}$ belongs to a particular class of matrices with dependencies in their entries. Such a class is formed by symmetric matrices, persymmetric matrices, skew-symmetric matrices, and others. Therefore, it is reasonable to consider subsets of S for which the elements x are solutions of linear systems $Ax=b$ with *special* matrices A only. This leads to the problem discussed in this paper. Our results are formulated in terms of inequalities involving the bounds of $[A]$, $[b]$. They can easily be reformulated using the midpoints $\check{A}$, $\check{b}$ and the tolerances ΔA, Δb, although a compact form such as (1.6) is still missing.

We also mention that the sets S_{sym} and S_{skew} were already considered in [8] and [9]. There, bounds for the projections of these sets onto the coordinate axes were derived but no characterization of these sets were given.

We have arranged our paper as follows. In section 2 we list the notation which we will use throughout the paper; in section 3 we present the results. We close our paper with some examples in section 4 which illustrate the technique and the theory.

2. Preliminaries. By $\boldsymbol{R}^n$, $\boldsymbol{R}^{n\times n}$, $\boldsymbol{IR}$, $\boldsymbol{IR}^n$, and $\boldsymbol{IR}^{n\times n}$ we denote the set of real vectors with n components, the set of real $n \times n$ matrices, the set of intervals, the set of interval vectors with n components, and the set of $n \times n$ interval matrices, respectively. By "interval" we always mean a real compact interval. Interval vectors and interval matrices are vectors and matrices, respectively, with interval entries. We write intervals in brackets with the exception of degenerate intervals (so-called *point intervals*), which we identify with the element being contained, and we proceed similarly with interval vectors and interval matrices. We write $[A] = [\underline{A}, \overline{A}] = ([a]_{ij}) = ([\underline{a}_{ij}, \overline{a}_{ij}]) \in \boldsymbol{IR}^{n\times n}$ simultaneously, without further reference, and we use an analogous notation for intervals and interval vectors. By $[A]^T$ we mean the transposed matrix of $[A]$. We mention that $[A] = [A]^T$ is equivalent to $\underline{A} = \underline{A}^T$ and $\overline{A} = \overline{A}^T$ and that $[A] = -[A]^T$ is equivalent to $\underline{A} = -\overline{A}^T$ and $\overline{A} = -\underline{A}^T$. Therefore, if an interval matrix $[A]$ fulfills the condition $[A] = -[A]^T$, its midpoint matrix $\check{A} := \frac{1}{2}(\underline{A} + \overline{A})$ satisfies $\check{A} = -\check{A}^T$; i.e., $\check{A}$ is skew-symmetric. We call an $n \times n$ interval matrix *singular* if it contains at least one singular real matrix; otherwise, we call it *regular*. For computations with interval quantities we refer to [1] or [10].

By O we denote any closed orthant of $\boldsymbol{R}^n$. To distinguish among the sets S, S_{sym}, S_{per}, and S_{skew} we call S_{sym} the *symmetric solution set*, S_{per} the *persymmetric solution set*, and S_{skew} the *skew-symmetric solution set*.

3. Results. We start this section with a topological result which for S and S_{sym} is already known (see [4]).

THEOREM 3.1. *Let $[A] \in \boldsymbol{IR}^{n\times n}$ be regular and satisfy* (1.5).

(a) *Each of the sets S_{sym}, S_{per}, S_{skew}, $S \cap O$, $S_{sym} \cap O$, $S_{per} \cap O$, and $S_{skew} \cap O$ is compact.*

(b) *Each of the sets S, S_{sym}, S_{per}, S_{skew}, and $S \cap O$ is connected; $S \cap O$ is convex.*

Proof. First, we prove the assertions for S_{skew}. Let $A = -A^T \in [A]$ and interpret $x = A^{-1}b$ as a function f of the $\frac{n(n-1)}{2}$ variables a_{ij}, $1 \le i < j \le n$ and the n variables b_i, $1 \le i \le n$. This function is continuous. Since $[a]_{ij}$, $[b]_i$ are connected and compact the same holds for the range S_{skew} of f.

The compactness of the intersection $S_{skew} \cap O$ follows from S_{skew} being compact and from O being closed.

In the cases of S, S_{sym}, and S_{per} one proves the assertions by similar arguments.

The convexity of $S \cap O$ results from the fact that this set can be expressed as the intersection of finitely many half spaces (cf. [11] or the subsequent Theorem 3.4, e.g.). □

Remark. If $[A]$ is singular but contains no singular *symmetric* matrix the proof of Theorem 3.1 shows that S_{sym} remains compact and connected and that $S_{sym} \cap O$ remains compact. An analogous statement holds for S_{per}, S_{skew}, $S_{per} \cap O$, and $S_{skew} \cap O$. For singular $[A]$ the solution set S, however, is empty or unbounded since the kernel of each singular matrix $A \in [A]$ is unbounded. Due to singularity, the function f with $f(A, b) := A^{-1}b$ is certainly not defined on $[A] \times [b]$. This already indicates that the assertions of Theorem 3.1 may be wrong in the singular case. As an illustration we consider the example

$$[A] := \begin{pmatrix} 0 & [-1,1] \\ [-1,1] & 0 \end{pmatrix}, \qquad [b] := \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Any real matrix $A \in [A]$ can be represented by

$$A = \begin{pmatrix} 0 & \alpha \\ \beta & 0 \end{pmatrix}$$

with α, $\beta \in [-1, 1]$. Hence A is regular with

$$A^{-1} = \begin{pmatrix} 0 & \beta^{-1} \\ \alpha^{-1} & 0 \end{pmatrix}$$

provided that $\alpha\beta \neq 0$. We obtain

$$S = \{ (\gamma, \delta)^T \mid \gamma \in \boldsymbol{R},\ -\infty < \delta \le -1 \text{ or } 1 \le \delta < \infty \},$$
$$S_{sym} = S_{skew} = \{ (0, \delta)^T \mid -\infty < \delta \le -1 \text{ or } 1 \le \delta < \infty \},$$

which shows that neither S nor S_{sym} nor S_{skew} is compact or connected in this case. □

Our next theorem characterizes S_{skew} by a set of inequalities. Its proof starts with

$$x \in S_{skew} \Longleftrightarrow \underline{b} \le Ax \le \overline{b},\ A = -A^T \in [A], \tag{3.1}$$

transforms the inequalities in a suitable way by introducing new variables z_{ij}, and continues by applying the Fourier–Motzkin elimination (see [14], e.g.) to replace the entries of A by their bounds $\underline{a}_{ij}$ and $\overline{a}_{ij}$, respectively.

THEOREM 3.2. *Let $[A] = -[A]^T \in \boldsymbol{IR}^{n \times n}$ with $[a]_{ii} = 0$, $i = 1, \ldots, n$, and let $[b] \in \boldsymbol{IR}^n$. Then for any orthant $O \subseteq \boldsymbol{R}^n$ the set $S_{skew} \cap O$ can be represented as an intersection of finitely many closed sets, the boundaries of which are quadrics or hyperplanes. The inequalities characterizing these hyperplanes and quadrics can be derived from $\underline{b} \le Ax \le \overline{b}$, $A = -A^T \in [A]$, $x \in O$ by means of the Fourier–Motzkin elimination.*

Proof. Step 1. Let (3.1) hold, fix an orthant O, and define

$$a_{ij}^- := \begin{cases} \underline{a}_{ij} & \text{if } x_i x_j \ge 0, \\ \overline{a}_{ij} & \text{if } x_i x_j < 0, \end{cases} \qquad a_{ij}^+ := \begin{cases} \overline{a}_{ij} & \text{if } x_i x_j \ge 0, \\ \underline{a}_{ij} & \text{if } x_i x_j < 0, \end{cases}$$
$$b_i^- := \begin{cases} \underline{b}_i & \text{if } x_i \ge 0, \\ \overline{b}_i & \text{if } x_i < 0, \end{cases} \qquad b_i^+ := \begin{cases} \overline{b}_i & \text{if } x_i \ge 0, \\ \underline{b}_i & \text{if } x_i < 0. \end{cases} \tag{3.2}$$

Note that the values of a_{ij}^-, a_{ij}^+, b_i^-, b_i^+ are constant as long as x remains in the same orthant and that they satisfy $a_{ij}^- = -a_{ji}^+$ and $a_{ii}^- = a_{ii}^+ = 0$. We first will see that (3.1) is equivalent to

$$\begin{cases} x \in S \quad \wedge \quad \exists\, z_{ij} \in \boldsymbol{R} \text{ such that} \\ \qquad \begin{cases} a_{ij}^- x_i x_j \le z_{ij} \le a_{ij}^+ x_i x_j, & i, j = 1, \ldots, n,\ i < j, \\ z_{ij} = -z_{ji}, & i, j = 1, \ldots, n, \\ b_i^- x_i \le \sum_{j=1}^n z_{ij} \le b_i^+ x_i, & i = 1, \ldots, n. \end{cases} \end{cases} \tag{3.3}$$

Setting $z_{ij} := a_{ij} x_i x_j$ immediately shows that "(3.1) ⇒ (3.3)." To prove the converse we will construct $A \in \boldsymbol{R}^{n \times n}$ such that $A = -A^T \in [A]$ and $Ax \in [b]$. Consider a fixed index pair i_0, j_0 and define $a_{i_0 j_0}$ according to the following procedure.

Case 1: $x_{i_0} = 0$. Since $x \in S$ by (3.3), there are real numbers $a^*_{i_0 j}$ for $j = 1, \dots, n$ such that

$$\underline{a}_{i_0 j} \le a^*_{i_0 j} \le \overline{a}_{i_0 j} \tag{3.4}$$

and

$$\underline{b}_{i_0} \le \sum_{j=1}^{n} a^*_{i_0 j} x_j \le \overline{b}_{i_0}. \tag{3.5}$$

If $x_{j_0} \ne 0$ then $a_{i_0 j_0} := a^*_{i_0 j_0} =: -a_{j_0 i_0}$; if $x_{j_0} = 0$ then $a_{i_0 j_0} := \check{a}_{i_0 j_0}$ with $\check{a}_{i_0 j_0}$ being the corresponding entry of the skew-symmetric midpoint matrix $\check{A} \in [A]$.

Case 2: $x_{i_0} \ne 0$. If $x_{j_0} \ne 0$ then $a_{i_0 j_0} := \frac{z_{i_0 j_0}}{x_{i_0} x_{j_0}}$; if $x_{j_0} = 0$ then $a_{i_0 j_0}$ is already defined by the preceding case when the roles of i_0 and j_0 are exchanged.

If one lets i_0 run from 1 to n and if for each fixed i_0 the second index in $z_{i_0 j_0}$ runs from 1 to n then by the procedure above a skew-symmetric matrix $A \in [A]$ is constructed which satisfies (3.1). Note that in Case 1 of our procedure there may occur several choices for the entries $a^*_{i_0 j}$ such that (3.4) and (3.5) are valid. It is obvious that in this case for a fixed i_0 the entries of one and the same double inequality (3.5) must be chosen for those $j_0 = 1, \dots, n$ for which $x_{j_0} \ne 0$. Together with the last double inequality in (3.3), this guarantees $\underline{b}_i \le \sum_{j=1}^{n} a_{ij} x_j \le \overline{b}_i$.

The condition "$x \in S$" in (3.3) is necessary, as the example $A := 0 \in \boldsymbol{R}^{1\times 1}$, $b := 1 \in \boldsymbol{R}$ shows. Here, $x = 0 \in \boldsymbol{R}$ is clearly not in $S \supseteq S_{skew}$, but the remaining conditions of (3.3) are fulfilled for $z_{11} = 0$.

Step 2. By $z_{ii} = -z_{ii}$ we obtain $z_{ii} = 0$. Therefore, we omit z_{ii} in (3.3). We now apply the Fourier–Motzkin elimination to (3.3). We illustrate this process by eliminating z_{12}. To this end we replace z_{ij} by $-z_{ji}$ for all $i > j$ in the inequalities of (3.3). We rewrite these inequalities and change their order by forming three groups: the inequalities of the first group have the form $\cdots \le z_{12}$ with z_{12}-free left-hand side, the inequalities of the second group read $z_{12} \le \cdots$ with z_{12}-free right-hand side, and the inequalities of the third group do not contain z_{12}. Since the maximum over all left-hand sides of the inequalities of the first group is less than or equal to the minimum over all right-hand sides of the inequalities of the second group, these inequalities are equivalent to requiring that *each* left-hand side of the first group be less than or equal to *each* right-hand side of the second group while keeping all inequalities of the third group. Omitting trivial inequalities, (3.3) is equivalent to

$$\left\{ \begin{array}{l} x \in S \quad \wedge \quad \exists\, z_{ij} \in \boldsymbol{R} \text{ such that} \\ \qquad \left\{ \begin{array}{l} a^-_{12} x_1 x_2 \le b^+_1 x_1 - \sum_{j=3}^{n} z_{1j}, \\ a^-_{12} x_1 x_2 \le -b^-_2 x_2 + \sum_{j=3}^{n} z_{2j}, \\ b^-_1 x_1 - \sum_{j=3}^{n} z_{1j} \le a^+_{12} x_1 x_2, \\ b^-_1 x_1 - \sum_{j=3}^{n} z_{1j} \le -b^-_2 x_2 + \sum_{j=3}^{n} z_{2j}, \\ -b^+_2 x_2 + \sum_{j=3}^{n} z_{2j} \le a^+_{12} x_1 x_2, \\ -b^+_2 x_2 + \sum_{j=3}^{n} z_{2j} \le b^+_1 x_1 - \sum_{j=3}^{n} z_{1j}, \\ \text{remaining (in)equalities of (3.3)}, \end{array} \right. \end{array} \right. \tag{3.6}$$

where z_{12} and z_{21} no longer occur. This process of eliminating z_{ij} can be continued until we end up with a set of final inequalities which (together with $x \in S \cap O$) is equivalent to $x \in S_{skew} \cap O$ and which contains no variable z_{ij}. This proves the theorem. □

At the end of the elimination process, there are two special inequalities for each $i \in \{1, \ldots, n\}$ which can be divided by $x_i \neq 0$ such that no fractions occur. For example, if the first inequality of (3.6) is combined successively with the inequalities $a_{1j}^- x_1 x_j \le z_{1j}$ one obtains the final inequality $\sum_{j=2}^n a_{1j}^- x_1 x_j \le b_1^+ x_1$. Since $a_{11}^- = a_{11}^+ = 0$ it can be supplemented to $\sum_{j=1}^n a_{1j}^- x_1 x_j \le b_1^+ x_1$, which reduces to

$$\sum_{j=1}^n a_{1j}^- x_j \le \overline{b}_1 \quad \text{if } x_1 > 0 \quad \text{and} \quad \sum_{j=1}^n a_{1j}^- x_j \ge \underline{b}_1 \quad \text{if } x_1 < 0. \tag{3.7}$$

From the third inequality of (3.6) one similarly obtains

$$\sum_{j=1}^n a_{1j}^+ x_j \ge \underline{b}_1 \quad \text{if } x_1 > 0 \quad \text{and} \quad \sum_{j=1}^n a_{1j}^+ x_j \le \overline{b}_1 \quad \text{if } x_1 < 0. \tag{3.8}$$

With

$$\hat{a}_{ij}^- := \begin{cases} \underline{a}_{ij} & \text{if } x_j \ge 0, \\ \overline{a}_{ij} & \text{if } x_j < 0, \end{cases} \qquad \hat{a}_{ij}^+ := \begin{cases} \overline{a}_{ij} & \text{if } x_j \ge 0, \\ \underline{a}_{ij} & \text{if } x_j < 0, \end{cases} \tag{3.9}$$

the four inequalities in (3.7) and (3.8) can be summarized to

$$\sum_{j=1}^n \hat{a}_{1j}^- x_j \le \overline{b}_1 \quad \text{and} \quad \sum_{j=1}^n \hat{a}_{1j}^+ x_j \ge \underline{b}_1,$$

provided that $x_1 \neq 0$. Repeating the arguments, one finally gets

$$\left. \begin{array}{l} \sum_{j=1}^n \hat{a}_{ij}^- x_j \le \overline{b}_i, \\ \sum_{j=1}^n \hat{a}_{ij}^+ x_j \ge \underline{b}_i, \end{array} \right\} \quad i = 1, \ldots, n \tag{3.10}$$

if no component of x equals 0. These inequalities are just those which characterize S and which are known as the Oettli–Prager theorem (cf. [11]), which we restate as Theorem 3.4. They can either be omitted in the list of inequalities if "$x \in S$" remains there as in (3.6), or "$x \in S$" can be cancelled when (3.10) is used instead. This last remark also holds if some of the components of x are zero.

We also note that the number $n_\#$ of final inequalities for $S_{skew} \cap O$ seems to be double exponential. Thus we could show that $n_\#$ is roughly bounded by $8 \cdot \left(\frac{3}{2}\right)^{2^{\kappa+1}}$ with $\kappa := \frac{n(n+1)}{2}$. Since the arguments are a little bit clumsy and the proof is lengthy we will skip it.

The same technique for S_{skew} can also be applied to construct a set of inequalities which characterize S_{sym} provided that $[A] = [A]^T$. To get the equivalence to "$x \in S_{sym}$" one must replace the equality in (3.1) by $A = A^T$, and one uses $z_{ij} = z_{ji}$ in (3.3) instead of $z_{ij} = -z_{ji}$. Analogously to Theorem 3.2, we get the following theorem.

THEOREM 3.3. *Let $[A] = [A]^T \in \boldsymbol{IR}^{n \times n}$ and let $[b] \in \boldsymbol{IR}^n$. Then for any orthant $O \subseteq \boldsymbol{R}^n$ the set $S_{sym} \cap O$ can be represented as an intersection of finitely many closed sets, the boundaries of which are quadrics or hyperplanes. The inequalities characterizing these hyperplanes and quadrics can be derived from the elimination process described above or they are of the form $x_i = 0$.* □

Theorem 3.3 can analogously be formulated for S_{per} since $Ax = b \Longleftrightarrow EAx = Eb$, whence S_{per} for A equals S_{sym} for $EA = (EA)^T$.

The solution set for other classes of special matrices such as Hankel or Toeplitz matrices shows particularities which essentially differ from those which we have presented up to now. Thus, the inequalities need no longer remain the same in a fixed orthant and they cannot be treated by means of the particular variables z_{ij}. Work in this respect is in progress.

Inequalities (3.10) can also be obtained with the technique above if one starts with

$$x \in S \tag{3.11}$$

instead of $x \in S_{skew}$. The conditions corresponding to (3.3) then read

$$\exists\, z_{ij} \in \boldsymbol{R} \text{ such that } \begin{cases} \hat{a}_{ij}^- x_j \le z_{ij} \le \hat{a}_{ij}^+ x_j, & i,j = 1,\dots,n, \\ \underline{b}_i \le \sum_{j=1}^n z_{ij} \le \overline{b}_i, & i = 1,\dots,n \end{cases} \tag{3.12}$$

with $\hat{a}_{ij}^-$, $\hat{a}_{ij}^+$ from (3.9). To prove the implication "(3.12) $\Rightarrow$ (3.11)" set $a_{ij} = \frac{z_{ij}}{x_j}$ if $x_j \neq 0$. If $x_j = 0$ then any element from $[a]_{ij}$ can be used to construct a matrix A such that $Ax \in [b]$ holds. It is easy to see that one ends up with inequalities (3.10) if one performs the elimination process as above, starting with (3.12).

For completeness we state the result in a separate theorem.

THEOREM 3.4 (Oettli–Prager theorem [11]). *Let $[A] \in \boldsymbol{IR}^{n\times n}$ and let $[b] \in \boldsymbol{IR}^n$. Then for any orthant $O \subseteq \boldsymbol{R}^n$ the set $S \cap O$ can be represented as the intersection of closed half spaces. These half spaces are given by*

$$\left.\begin{aligned} \textstyle\sum_{j=1}^n \hat{a}_{ij}^- x_j \le \overline{b}_i, \\ \textstyle\sum_{j=1}^n \hat{a}_{ij}^+ x_j \ge \underline{b}_i, \end{aligned}\right\} \quad i = 1,\dots,n \tag{3.13}$$

or

$$x_i \le 0 \quad \text{or} \quad x_i \ge 0, \tag{3.14}$$

where the inequalities in (3.14) are used to characterize the orthant O and where $\hat{a}_{ij}^-$, $\hat{a}_{ij}^+$ are defined in (3.9). $\square$

4. Examples. In this section we present several examples to illustrate the results of section 3. In particular, we construct the inequalities for characterizing S, S_{sym}, S_{per}, and S_{skew}.

In our first example we consider 2×2 interval matrices.

Example 4.1.

(a) Let $[A] \in \boldsymbol{IR}^{2\times 2}$, $[b] \in \boldsymbol{IR}^2$. Then S is characterized according to (3.13) by the inequalities

$$\begin{cases} \hat{a}_{11}^- x_1 + \hat{a}_{12}^- x_2 \le \overline{b}_1, & \hat{a}_{11}^+ x_1 + \hat{a}_{12}^+ x_2 \ge \underline{b}_1, \\ \hat{a}_{21}^- x_1 + \hat{a}_{22}^- x_2 \le \overline{b}_2, & \hat{a}_{21}^+ x_1 + \hat{a}_{22}^+ x_2 \ge \underline{b}_2 \end{cases} \tag{4.1}$$

with the coefficients according to (3.9).

(b) Let $[A] = [A]^T$ hold. The symmetric solution set S_{sym} is described by the four inequalities in (4.1) supplemented by the two inequalities

$$\begin{cases} b_1^- x_1 - b_2^+ x_2 - a_{11}^+ x_1^2 + a_{22}^- x_2^2 \le 0, \\ -b_1^+ x_1 + b_2^- x_2 + a_{11}^- x_1^2 - a_{22}^+ x_2^2 \le 0 \end{cases} \tag{4.2}$$

with the coefficients from (3.2). These inequalities show that the boundary of S_{sym} can already be curvilinear in the 2×2 case.

(c) Let $E[A] = (E[A])^T$ hold. The persymmetric solution set S_{per} is described by the four inequalities in (4.1) supplemented by the two inequalities in (4.2) if one redefines $a_{ii}^{\pm}$, $b_i^{\pm}$ appropriately.

(d) Let $[A] = -[A]^T$ hold with $[a]_{ii} = 0$ for $i = 1, 2$. The skew-symmetric solution set S_{skew} is given by the four inequalities in (4.1) with $\hat{a}_{ii}^- = \hat{a}_{ii}^+ = 0$ in addition to the two inequalities

$$b_1^- x_1 \le -b_2^- x_2, \qquad -b_2^+ x_2 \le b_1^+ x_1, \tag{4.3}$$

which follow directly from (3.6) taking into account $z_{11} = z_{22} = 0$. The skew-symmetric solution set in $\boldsymbol{R}^2$ is apparently bounded by a polygon; i.e., its boundary is formed by straight lines. Taking into account $\hat{a}_{ii}^- = \hat{a}_{ii}^+ = 0$, one sees immediately from (4.1) that the solution set S is an interval vector. This is not always the case for S_{skew}. For example, choose $[b] := (1,1)^T$ and $[a]_{12} := [0.25, 1]$. Then any skew-symmetric element A of $[A]$ can be written in the form

$$A = \alpha \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = -\alpha^2 A^{-1} \qquad \text{with } 0.25 \le \alpha \le 1.$$

Hence $S_{skew} = \{\, \beta(-1,1)^T \mid 1 \le \beta \le 4 \}$; i.e., S_{skew} is the straight line in the plane between the points $(-1, 1)$ and $(-4, 4)$. The corresponding solution set S, however, is given by

$$S = \{\, (-\beta, \gamma)^T \mid 1 \le \beta,\, \gamma \le 4 \,\} = ([-4, -1], [1, 4])^T. \qquad \square$$

In our second example we consider 3×3 tridiagonal interval matrices.

Example 4.2.

(a) Let $[A] \in \boldsymbol{IR}^{3\times 3}$ with $[a]_{13} = [a]_{31} := 0$, and let $[b] \in \boldsymbol{R}^3$. Then S is characterized by the inequalities

$$\left\{ \begin{array}{rr} \hat{a}_{11}^- x_1 + \hat{a}_{12}^- x_2 \le \bar{b}_1, & \hat{a}_{11}^+ x_1 + \hat{a}_{12}^+ x_2 \ge \underline{b}_1, \\ \hat{a}_{21}^- x_1 + \hat{a}_{22}^- x_2 + \hat{a}_{23}^- x_3 \le \bar{b}_2, & \hat{a}_{21}^+ x_1 + \hat{a}_{22}^+ x_2 + \hat{a}_{23}^+ x_3 \ge \underline{b}_2, \\ \hat{a}_{32}^- x_2 + \hat{a}_{33}^- x_3 \le \bar{b}_3, & \hat{a}_{32}^+ x_2 + \hat{a}_{33}^+ x_3 \ge \underline{b}_3, \end{array} \right. \tag{4.4}$$

where the coefficients are again given by (3.9).

(b) For tridiagonal 3×3 matrices $[A] = [A]^T$ the symmetric solution set S_{sym} is characterized by the six inequalities in (4.4) and by the four additional inequalities

$$\left\{ \begin{array}{r} +b_1^- x_1 - b_2^+ x_2 - a_{11}^+ x_1^2 + a_{22}^- x_2^2 + a_{23}^- x_2 x_3 \le 0, \\ +b_1^- x_1 - b_2^+ x_2 + b_3^- x_3 - a_{11}^+ x_1^2 + a_{22}^- x_2^2 - a_{33}^+ x_3^2 \le 0, \\ +b_1^- x_1 - (+b_2^+ - b_2^-) x_2 - a_{11}^+ x_1^2 - a_{12}^+ x_1 x_2 - (+a_{22}^+ - a_{22}^-) x_2^2 \le 0, \\ +b_2^- x_2 - b_3^+ x_3 - a_{12}^+ x_1 x_2 - a_{22}^+ x_2^2 + a_{33}^- x_3^2 \le 0 \end{array} \right. \tag{4.5}$$

together with their four counterparts, which one gets by replacing each minus sign by a plus sign, and vice versa (also in the superscripts). The coefficients of (4.5) are defined in (3.2). Note that the information of the third inequality in (4.5) is contained in that of the first row of (4.4) if $[b]_2$ and $[a]_{22}$ are point intervals.

Without proof we mention that the number of inequalities for S_{sym} increases to 44 for a dense 3×3 system.

(c) The skew-symmetric solution set S_{skew} is characterized by (4.4) with $\hat{a}_{ii}^- = \hat{a}_{ii}^+ = 0$ for $i = 1, 2, 3$ and by the inequalities

$$
(4.6) \qquad \left\{ \begin{array}{r} -b_1^+ x_1 - b_2^+ x_2 + a_{23}^- x_2 x_3 \leq 0, \\ +b_1^- x_1 + b_2^- x_2 + b_3^- x_3 \leq 0, \\ +b_1^- x_1 - (+b_2^+ - b_2^-) x_2 - a_{12}^+ x_1 x_2 \leq 0, \\ -b_2^+ x_2 - b_3^+ x_3 - a_{12}^+ x_1 x_2 \leq 0 \end{array} \right.
$$

together with their four counterparts, which are defined analogously as for S_{sym}. The inequalities in (4.6) look similar to those in (4.5) when taking into account $[a]_{ii} = 0$ for $i = 1, 2, 3$. Again, the third inequality in (4.6) equals the first one in (4.4) if $[b_2]$ is a point interval. Note also that according to section 1 each skew-symmetric matrix from $\boldsymbol{R}^{3\times 3}$ is singular! $\square$

In our third example we describe S and S_{skew} in two different ways, a direct way (feasible since there is only one nontrivial pair of intervals) and a second way where we will apply the results of Example 4.2.

Example 4.3. Let

$$
[A] := \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & [0.5, 1] \\ 0 & [-1, -0.5] & 0 \end{pmatrix}, \qquad [b] := \begin{pmatrix} [0, 2] \\ 0 \\ -1 \end{pmatrix}.
$$

Then $[A] = -[A]^T$ with $[a]_{ii} = 0$, $i = 1, 2, 3$. Each $A \in [A]$, $b \in [b]$ can be represented as

$$
A = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & \alpha \\ 0 & -\beta & 0 \end{pmatrix}, \qquad b = \begin{pmatrix} \gamma \\ 0 \\ -1 \end{pmatrix}
$$

with α, $\beta \in [0.5, 1]$, $\gamma \in [0, 2]$. The linear system $Ax = b$ then reads

$$
(4.7) \qquad x_2 = \gamma,
$$

$$
(4.8) \qquad -x_1 + \alpha x_3 = 0,
$$

$$
(4.9) \qquad -\beta x_2 = -1.
$$

(a) We first want to describe the solution set S. Equations (4.7) and (4.8) show that $x_2 \geq 0$ and $\mathrm{sign}(x_1 x_3) \geq 0$. This means that only the first orthant O_1 and the sixth orthant O_6 can contain elements of S, where O_1 is characterized by $x_i \geq 0$, $i = 1, 2, 3$, and where O_6 is given by $x_1 \leq 0$, $x_2 \geq 0$, $x_3 \leq 0$. By the first and the third equation the system (4.7)–(4.9) is solvable if and only if $\beta\gamma = 1$. This is possible for any $\beta \in [0.5, 1]$ since $\gamma = \beta^{-1} \in [1, 2] \subseteq [0, 2]$. The solution can be rewritten as

$$
(4.10) \qquad x_1 = \alpha x_3, \quad x_2 = \beta^{-1}, \quad x_3 \in \boldsymbol{R}.
$$

For each fixed α, $\beta \in [0.5, 1]$ these equations represent, of course, a straight line which lies in the plane $x_2 = \beta^{-1} \in [1, 2]$ and which crosses the x_2-axis at $(0, \beta^{-1}, 0)$. For each fixed $\beta \in [0.5, 1]$ one thus gets a (double) sector in $O_1 \cup O_6$ which is bounded by the straight lines $x_1 = 0.5x_3$ and $x_1 = x_3$ while $x_2 = \beta^{-1}$. Varying β results in two wedges, the cutting edges of which have length 1 and meet at the x_2-axis from $(0, 1, 0)$ to $(0, 2, 0)$.

(b) To characterize S_{skew} let $\alpha = \beta$. From (4.10) we then obtain $x_1 x_2 = x_3$ with $x_2 \in [1, 2]$, i.e., S_{skew} is the intersection of S with the hyperbolical paraboloid

$x_3 = x_1x_2$ which transforms to $y_3 = y_1^2 - y_2^2$ via $x_1 = y_1 + y_2$, $x_2 = y_1 - y_2$, $x_3 = y_3$. In particular, the boundary of S_{skew} is curvilinear. Figure 1 shows $S \cap O_1$ and $S_{skew} \cap O_1$. The intersections $S \cap O_6$ and $S_{skew} \cap O_6$ are obtained by rotating the two sets around the x_2-axis by an amount of 180° degrees.

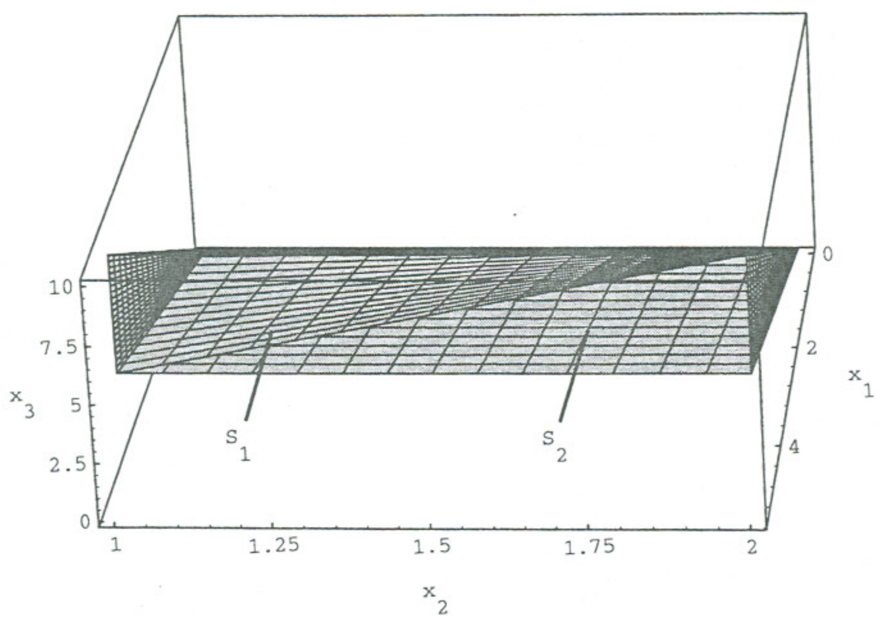

FIG. 4.1. *The shape of the solution sets* $S_1 := S \cap O_1$, $S_2 := S_{skew} \cap O_1$ *in Example* 4.3.

(c) We now want to describe S and S_{skew} in a second way, namely, by the inequalities resulting from (4.4) and (4.6). For simplicity we use $S \subseteq O_1 \cup O_6$, which yields $a_{23}^- = 0.5 = -a_{32}^+$, $a_{23}^+ = 1 = -a_{32}^-$. Inequalities (4.4) can then be written in the form

$$0 \le x_2 \le 2, \tag{4.11}$$

$$0.5x_3 \le x_1 \le x_3, \tag{4.12}$$

$$1 \le x_2 \le 2 \tag{4.13}$$

if $(x_1, x_2, x_3) \in O_1$. In O_6 inequality (4.12) must be replaced by $x_3 \le x_1 \le 0.5x_3$. Since (4.13) is more restrictive than (4.11) we can omit (4.11). Thus S is characterized by (4.12) and (4.13).

Inequalities (4.6) and their counterparts yield to

$$b_1^- x_1 \le x_2x_3 \le 2b_1^+ x_1, \tag{4.14}$$

$$b_1^- x_1 \le x_3 \le b_1^+ x_1, \tag{4.15}$$

$$b_1^- x_1 \le x_1x_2 \le b_1^+ x_1, \tag{4.16}$$

$$x_3 = x_1x_2 \tag{4.17}$$

in O_1; in O_6 inequality (4.14) must be exchanged by $2b_1^- x_1 \le x_2 x_3 \le b_1^+ x_1$. Dividing (4.16) by x_1 implies (4.11). Hence (4.16) can be omitted. Since (4.15) is identical with (4.16) if (4.17) is used, we can skip (4.15) too. Replacing x_3 in (4.14) by (4.17) and dividing by x_1 yields to $0 = \underline{b}_1 \le x_2^2 \le 2\overline{b}_1 \le 4$, which again is fulfilled if (4.11) holds. Therefore, the inequalities for S_{skew} reduce in O_1 to

$$\begin{aligned} 1 &\le x_2 \le 2, \\ x_1 &\le x_3 \le 2x_1, \\ x_3 &= x_1 x_2, \end{aligned}$$

which is equivalent to (4.10) when taking into account $\alpha = \beta \in [0.5, 1]$. The same holds in O_6 if the second double inequality is replaced by $2x_1 \le x_3 \le x_1$. □

In our last example we consider a 2×2 interval matrix $[A]$ which satisfies $[A] = [A]^T$.

Example 4.4. Let

$$[A] := \begin{pmatrix} 1 & [0,1] \\ [0,1] & [-4,-1] \end{pmatrix}, \qquad [b] := \begin{pmatrix} [0,2] \\ [0,2] \end{pmatrix}.$$

Then $[A] = [A]^T$ with

$$A = \begin{pmatrix} 1 & \alpha \\ \beta & -\gamma \end{pmatrix} \in [A] \quad \Longrightarrow \quad A^{-1} = \frac{1}{\gamma + \alpha\beta} \begin{pmatrix} \gamma & \alpha \\ \beta & -1 \end{pmatrix}$$

with $\alpha, \beta \in [0,1]$, $\gamma \in [1,4]$. Since $\underline{b} \ge 0$ the first component of $A^{-1}b$ is nonnegative for all $b \in [b]$. Therefore, S is completely contained in the union $O_1 \cup O_4$ of the first and the fourth quadrants.

We first consider $S \cap O_1$. According to (4.1) we get the inequalities

$$x_1 \le 2, \quad x_2 \ge -0.5, \quad x_2 \ge -x_1, \quad x_1 \ge x_2. \tag{4.18}$$

This means that $S \cap O_1$ is the triangle with the corners $(0,0)$, $(2,0)$, and $(2,2)$.

The corresponding inequalities for $S \cap O_4$ are given by

$$x_1 \ge 0, \quad x_2 \ge -2, \quad x_2 \le 2 - x_1, \quad x_2 \le 0.25\, x_1. \tag{4.19}$$

They describe a quadrangle with the corners $(0,0)$, $(0,-2)$, $(4,-2)$, and $(2,0)$.

To describe $S_{sym} \cap O_1$ we need inequalities (4.18) and the two inequalities from (4.2), which can be transform to

$$4x_1^2 + (4x_2 + 1)^2 \ge 1, \qquad (x_1 - 1)^2 + x_2^2 \le 1. \tag{4.20}$$

The first inequality of (4.20) describes an ellipse and its exterior. Since the ellipse lies completely in the lower half plane the first inequality of (4.20) is no restriction for $S_{sym} \cap O_1$. The second inequality describes a closed disc D_1 with center $(1,0)$ and radius 1. The boundary of the intersection with $S \cap O_1$ is formed by the straight line from $(0,0)$ to $(1,1)$, the part of the circle ∂D_1 from $(1,1)$ to $(2,0)$, and the part of the x_1-axis from $(2,0)$ back to $(0,0)$.

The inequalities in (4.19) together with the two inequalities

$$x_1^2 + 4x_2^2 \ge 0, \qquad (x_1 - 1)^2 + (x_2 + 1)^2 \le 2 \tag{4.21}$$

characterize $S_{sym} \cap O_4$. The first inequality in (4.21) is always true. The second inequality describes a disc D_2 with center $(1,-1)$ and radius $\sqrt{2}$. The boundary of its intersection with $S \cap O_4$ is formed by the straight lines from $(0,0)$ to $(0,-2)$, from $(0,-2)$ to $(2,-2)$, and from $(2,0)$ to $(0,0)$, and by the part of the circle ∂D_2 from $(2,-2)$ to $(2,0)$. The situation is illustrated by Figure 2. □

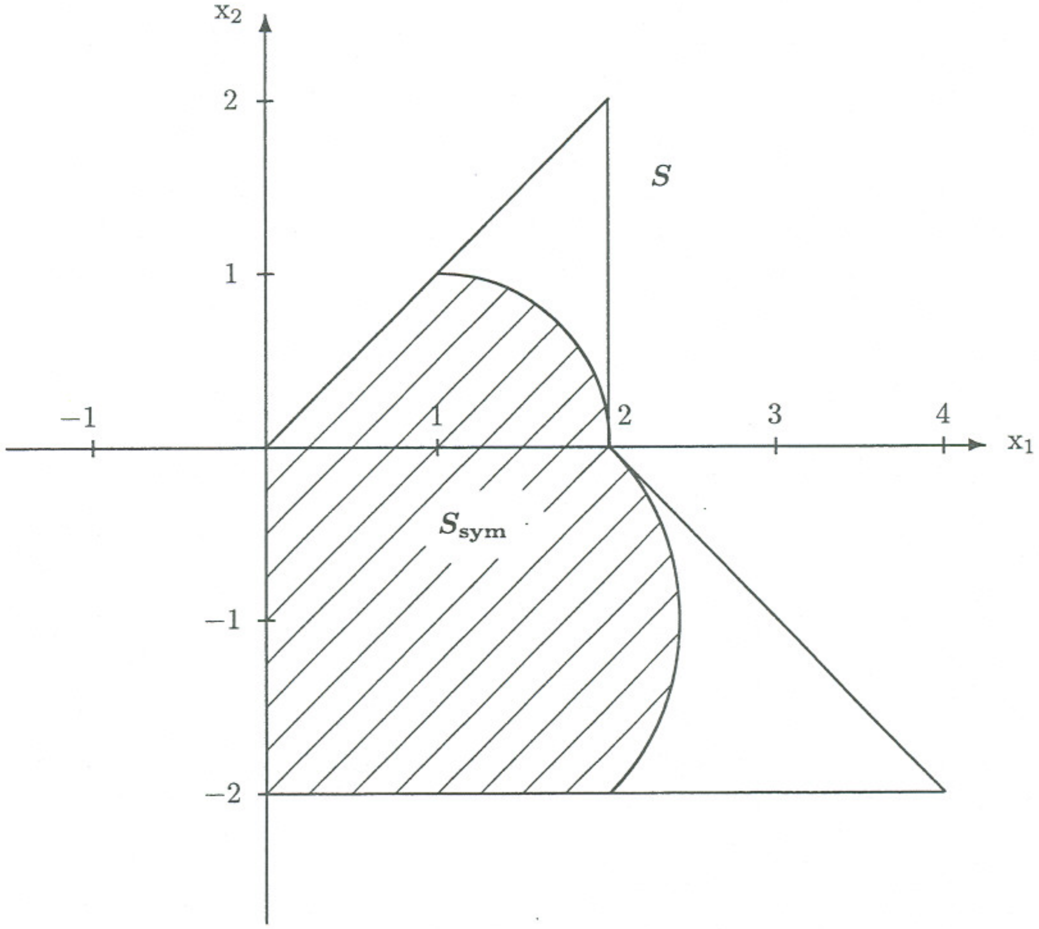

FIG. 4.2. *The shape of the solution sets S, S_{sym} in Example* 4.4.

Acknowledgment. We thank both referees for their valuable suggestions and remarks which improved this paper.

REFERENCES

[1] G. ALEFELD AND J. HERZBERGER, *Introduction to Interval Computations*, Academic Press, New York, 1983.

[2] G. ALEFELD, V. KREINOVICH, AND G. MAYER, *The shape of the symmetric solution set*, in Applications of Interval Computations, R. B. Kearfott and V. Kreinovich, eds., Kluwer, Boston, MA, 1995, pp. 61–79.

[3] G. ALEFELD, V. KREINOVICH, AND G. MAYER, *Symmetric linear systems with perturbed input data*, in Numerical Methods and Error Bounds, G. Alefeld and J. Herzberger, eds., Akademie Verlag, Berlin, 1996, pp. 16–22.

[4] G. ALEFELD AND G. MAYER, *On the symmetric and unsymmetric solution set of interval systems*, SIAM J. Matrix Anal. Appl., 16 (1995), pp. 1223–1240.

[5] H. BEECK, *Über Struktur und Abschätzungen der Lösungsmenge von linearen Gleichungssystemen mit Intervallkoeffizienten*, Computing, 10 (1972), pp. 231–244.

[6] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, MD, 1983.

[7] D. J. Hartfiel, *Concerning the solution set of $Ax = b$ where $P \leq A \leq Q$ and $p \leq b \leq q$*, Numer. Math., 35 (1980), pp. 355–359.
[8] C. Jansson, *Rigorous sensitivity analysis for real symmetric matrices with uncertain data*, in Computer Arithmetic, Scientific Computation and Mathematical Modelling, E. Kaucher, S. M. Markov, and G. Mayer, eds., Baltzer, Basel, Switzerland, 1991, pp. 293–316.
[9] C. Jansson, *Interval linear systems with symmetric matrices, skew-symmetric matrices and dependencies in the right hand side*, Computing, 46 (1991), pp. 265–274.
[10] A. Neumaier, *Interval Methods for Systems of Equations*, Cambridge University Press, Cambridge, UK, 1990.
[11] W. Oettli and W. Prager, *Compatibility of approximate solution of linear equations with given error bounds for coefficients and right-hand sides*, Numer. Math., 6 (1964), pp. 405–409.
[12] J. Rohn, *Interval linear systems*, Freiburger Intervall-Berichte, 84/7 (1984), pp. 33–58.
[13] S. M. Rump, *Verification methods for dense and sparse systems of equations*, in Topics in Validated Computations, J. Herzberger, ed., Elsevier, Amsterdam, The Netherlands, 1994, pp. 63–135.
[14] A. Schrijver, *Theory of Linear and Integer Programming*, John Wiley, New York, 1986.