On the Solution Sets of Particular Classes of Linear Interval Systems

Götz Alefeld*, Vladik Kreinovich† and Günter Mayer‡

Abstract

We characterize the solution set S of real linear systems $Ax = b$ by a set of inequalities if b lies between some given bounds $\underline{b}$, $\overline{b}$ and if the $n \times n$ coefficient matrix A varies similarly between two bounds $\underline{A}$ and $\overline{A}$. In addition, we restrict A to a particular class of matrices, for instance the class of the symmetric, the skew–symmetric, the persymmetric, the Toeplitz, and the Hankel matrices, respectively. In this way we generalize the famous Oettli–Prager criterion [7], results by Hartfiel [5] and the contents of the papers [1], [2].

Mathematics Subject Classification (1991): 65F05, 65G99

Keywords: Linear systems; solution set; interval matrix; Oettli–Prager criterion; Fourier–Motzkin elimination; symmetric matrices; Hankel matrices; Toeplitz matrices

1 Introduction

When solving $n \times n$ linear systems $Ax = b$ on a computer, the coefficients of the matrix A and the righthand side b are not always representable by machine numbers. Therefore one often solves linear systems $\tilde{A}x = \tilde{b}$ with input data $\tilde{A}$, $\tilde{b}$ which differ slightly from the original ones, i. e., with $\tilde{A}$ and $\tilde{b}$ from some interval quantities $[A]$ and [b], respectively, which also contain A, b. Sometimes one is also interested in the solutions of linear systems in which, in advance, the input data A and b are unknown to a certain extent. In this case, they normally are also limited to some $n \times n$ interval matrix $[A]$ and to some interval vector $[b]$ with n components. Therefore, it is an interesting question to discuss how the set

$$S := \{x \in \mathbb{R}^n | \ Ax = b, \ A \in [A], \ b \in [b]\} \tag{1}$$

looks like provided that $[A]$ does not contain a singular matrix. This question was answered in [5], [7], e. g., where it was shown that the intersection of S with any orthant O of $\mathbb{R}^n$ can be described by a set of linear inequalities which characterize a

*Institut für Angewandte Mathematik, Universität Karlsruhe, D–76128 Karlsruhe, Germany
†Department of Computer Science, University of Texas at El Paso, El Paso, TX 79968, USA
‡Fachbereich Mathematik, Universität Rostock, D–18051 Rostock, Germany

compact, convex polyhedron in O. The union of the corresponding polyhedrons of all orthants forms the set S which needs no longer to be convex but which remains a compact polyhedron and is therefore a connected set. This result was generalized in [1], [2], where only linear systems with *symmetric* matrices $A \in [A]$ were considered. It was shown that in each orthant O the corresponding set

$$S_{\mathrm{sym}} := \{x \in \mathbb{R}^n |\ Ax = b,\ A = A^T \in [A],\ b \in [b]\} \subseteq S \tag{2}$$

is the intersection of S with compact sets whose boundaries are quadrics, i. e., $S_{\mathrm{sym}} \cap O$ is described by a set of linear and quadratic inequalities. A similar result holds for the skew–symmetric matrices from $[A]$ and for the persymmetric ones, respectively, as was proved in [2].

A. Neumaier already drew attention to S_{sym} in a letter to J. Rohn in 1986 [8]. Bounds for S_{sym} can be obtained by methods in [4] and [6]; see also [9], where linear dependencies of the entries in A, b are allowed.

In [3] it was shown, that each projection of the set of linear systems $Ax = b$ with $A \in [A]$, $b \in [b]$ can be described by means of algebraic inequalities if the coefficients of A and b depend linearly on at most finitely many additional parameters, i.e.,

$$a_{ij} = a_{ij,0} + \sum_{\mu=1}^{m} a_{ij,\mu} u_\mu \quad \text{and} \quad b_i = b_{i,0} + \sum_{\mu=1}^{m} b_{i,\mu} u_\mu, \tag{3}$$

where $a_{ij,\mu}$, $b_{i,\mu}$, $\mu = 0, \ldots, m$, are real constants and where $u_\mu \in \mathbb{R}$, $\mu = 1, \ldots m$, are real parameters which vary in given compact intervals $[u]_\mu = [\underline{u}_\mu, \overline{u}_\mu]$. It was shown that even the converse holds, i. e., that every finite union of subsets each of which is described by algebraic inequalities can be represented as a projection of the solution set of linear equations $Ax = b$ of the above–mentioned form. This result was proved without presenting the constructive process explicitly which leads to the inequalities.

In the present paper, we will fill this gap. To this end we derive a central theorem in Section 3.1 which is basic for all the subsequent considerations and which ressembles the Fourier–Motzkin elimination (see [10], e. g.). It shows how parameters in a set of inequalities can be removed successively. This result can be applied to general matrices, to symmetric matrices, skew–symmetric matrices, persymmetric matrices, Hankel and Toeplitz matrices contained in a given interval matrix $[A]$ in order to characterize the corresponding solution set by a set of inequalities. For the symmetric, persymmetric, and skew–symmetric matrices the starting point differs now from that in [1], [2]; this time, it is more elementary. We also will outline the particularities which occur, when describing these solution sets. Thus, it is interesting to see that for particular solution sets the degree of the polynomials in the algebraic inequalities can be greater than two and that these inequalities seem to change in a fixed orthant O in contrast to the case $S \cap O$ and $S_{\mathrm{sym}} \cap O$. We will address to these problems in Section 3.

2 Notations

By $\mathbb{R}^n$, $\mathbb{R}^{n\times n}$, $\mathbb{IR}$, $\mathbb{IR}^n$, $\mathbb{IR}^{n\times n}$ 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. As usual, we denote the lower and upper bound of an interval $[a]$ by $\underline{a}$ and $\overline{a}$, respectively. Similarly, we write $[A] = [\underline{A}, \overline{A}] = ([a_{ij}]) = ([\underline{a}_{ij}, \overline{a}_{ij}]) \in \mathbb{IR}^{n\times n}$ without further reference. We call $[A] \in \mathbb{IR}^{n\times n}$ *regular* if it contains no singular matrix $A \in \mathbb{R}^{n\times n}$.

We denote any orthant of $\mathbb{R}^n$ by O and the first orthant by O_1. As usual, we call $A \in \mathbb{R}^{n\times n}$ *persymmetric* if $a_{ij} = a_{kl}$ for $k = n+1-j$, $l = n+1-i$, i.e., if it is symmetric with respect to the northeast–southwest diagonal, we call it a *Hankel matrix* if $a_{ij} = a_{kl}$ for $i+j = k+l$, i.e., if its entries are constant along each northeast – southwest diagonal, and a *Toeplitz matrix* if $a_{ij} = a_{kl}$ for $i-j = k-l$, i.e., if its entries are constant along each northwest–southeast diagonal for all indices $i, j, k, l \in \{1, \ldots, n\}$.

3 Results

3.1 A central theorem

We start this section with a theorem, which forms the basis for our subsequent considerations. It contains the constructive process which, for fixed x, is just the Fourier–Motzkin elimination (cf. [10], e.g.) and which leads to the inequalities mentioned in Section 1. In order to motivate the theorem we start by an example which shows how to describe S from (1) in a fixed orthant by means of inequalities as was done by Hartfiel in [5]. For simplicity we restrict ourselves to $S_1 := S \cap D$ where $D := O_1$. Trivially, S_1 is characterized by

$$S_1 = \{\, x \in D \mid \exists\, a_{ij}, b_i \in \mathbb{R} :\ (4) - (6) \text{ hold} \,\}$$

where

$$\sum_{i=1}^{n} a_{ij}x_j \le b_i \le \sum_{i=1}^{n} a_{ij}x_j, \quad i = 1, \ldots, n, \tag{4}$$

$$\underline{a}_{ij} \le a_{ij} \le \overline{a}_{ij}, \qquad i, j = 1, \ldots, n, \tag{5}$$

$$\underline{b}_i \le b_i \le \overline{b}_i, \qquad i = 1, \ldots, n\,. \tag{6}$$

Those inequalities in (4) – (5) which contain a_{11} can be rewritten as

$$b_1 - \sum_{i=2}^{n} a_{1j}x_j \ \le\ a_{11}x_1, \tag{7}$$

$$\underline{a}_{11} \ \le\ a_{11}, \tag{8}$$

$$a_{11}x_1 \;\le\; b_1 - \sum_{i=2}^{n} a_{1j}x_j, \qquad (9)$$

$$a_{11} \;\le\; \overline{a}_{11}. \qquad (10)$$

Multiply (8) and (10) by x_1 and combine each left–hand side of (7), (8) with each right–hand side of (9), (10) and drop the two trivial inequalities. Then this action results in the two nontrivial inequalities

$$b_1 - \sum_{i=2}^{n} a_{1j}x_j \;\le\; \overline{a}_{11}x_1, \qquad (11)$$

$$\underline{a}_{11}x_1 \;\le\; b_1 - \sum_{i=2}^{n} a_{1j}x_j, \qquad (12)$$

which are supplemented by

$$\text{the original } a_{11}\text{–free inequalities.} \qquad (13)$$

Hence

$$S_1 \subseteq S_2 := \{\, x \in D \mid \exists\, a_{ij}\ (i \ne 1 \text{ if } i = j), b_i \in \mathbb{R} :\ (11) - (13) \text{ hold}\,\}.$$

Since the converse $S_2 \subseteq S_1$ is also true (see the proof of the subsequent theorem) one ends up with $S_1 = S_2$, where in S_2 the entry a_{11} is replaced by the bounds $\underline{a}_{11}$, $\overline{a}_{11}$ of the given interval $[a]_{11}$. It is obvious that this process can be repeated for the remaining entries a_{ij} and b_i. One finally gets the inequalities in [5] which were derived there in a different manner.

We will generalize this elimination procedure in the subsequent theorem. There we do no longer distinguish between a_{ij} and b_i but introduce parameters u_μ, $\mu = 1,\dots,m$, instead. Moreover, we replace the constants 1 in front of b_i and a_{ij} in (5) and (6) (which we did not write down, of course) and the linear expressions x_j behind a_{ij} in (4) by more general expressions $f_{\lambda\mu}(x)$, $x \in D \subseteq \mathbb{R}^n$, and the constants $\underline{a}_{ij}$, $\overline{a}_{ij}$, $\underline{b}_i$, $\overline{b}_i$ by expressions $g_\lambda(x)$ which are independent on the parameters u_μ. For simplicity we also cancel the inequalities of the form (13) since they remain unchanged in S_1 as well as in S_2.

Theorem 1

Let $f_{\lambda\mu}$, g_λ, $\lambda = 1,\dots,k\ (\ge 2)$, $\mu = 1,\dots,m$, be real valued functions of $x = (x_1,\dots,x_n)^T$ on some subset $D \subseteq \mathbb{R}^n$. Assume that there is a positive integer $k_1 < k$ such that

$$f_{\lambda 1}(x) \not\equiv 0 \textit{ for all } \lambda \in \{1,\dots,k\}, \qquad (14)$$

$$f_{\lambda 1}(x) \ge 0 \textit{ for all } x \in D \textit{ and all } \lambda \in \{1,\dots,k\}, \qquad (15)$$

$$\textit{for each } x \in D \textit{ there is an index } \beta^* = \beta^*(x) \in \{1,\dots,k_1\} \textit{ with } f_{\beta^* 1}(x) > 0$$
$$\textit{and an index } \gamma^* = \gamma^*(x) \in \{k_1+1,\dots,k\} \textit{ with } f_{\gamma^* 1}(x) > 0. \qquad (16)$$

For m parameters $u_1,\dots,u_m$ varying in $\mathbb{R}$ and for x varying in D define the sets S_1, S_2 by

$$S_1 \;:=\; \{\, x \in D \mid \exists\, u_\mu \in \mathbb{R},\ \mu = 1,\dots,m :\ \textit{(17), (18) hold}\},$$
$$S_2 \;:=\; \{\, x \in D \mid \exists\, u_\mu \in \mathbb{R},\ \mu = 2,\dots,m :\ \textit{(19) holds}\},$$

where the inequalities (17), (18) and (19), respectively, are given by

$$g_\beta(x) + \sum_{\mu=2}^{m} f_{\beta\mu}(x)u_\mu \le f_{\beta 1}(x)u_1, \qquad \beta = 1, \ldots, k_1, \tag{17}$$

$$f_{\gamma 1}(x)u_1 \le g_\gamma(x) + \sum_{\mu=2}^{m} f_{\gamma\mu}(x)u_\mu, \qquad \gamma = k_1+1, \ldots, k\,; \tag{18}$$

and

$$g_\beta(x)f_{\gamma 1}(x) + \sum_{\mu=2}^{m} f_{\beta\mu}(x)f_{\gamma 1}(x)u_\mu \le g_\gamma(x)f_{\beta 1}(x) + \sum_{\mu=2}^{m} f_{\gamma\mu}(x)f_{\beta 1}(x)u_\mu, \tag{19}$$
$$\beta = 1, \ldots, k_1,\ \gamma = k_1+1, \ldots, k\,.$$

(Trivial inequalities such as $0 \le 0$ *can be omitted.)*

Then

$$S_1 = S_2.$$

Before proving Theorem 1 we remark that the parameter u_1 which occurs in the definition of S_1 is no longer needed in order to describe S_2. Therefore, we call the transition from the inequalities (17), (18) to the inequalities in (19) the *elimination of* u_1.

It is obvious that the assertion of Theorem 1 remains true if the inequalities in (17), (18) and the inequalities in (19) are supplemented by inequalities which do not contain the parameter u_1, as long as these inequalities are the same in both cases.

Proof of Theorem 1 .

$S_1 \subseteq S_2$:

W.l.o.g. let $S_1 \neq \emptyset$, fix $x \in S_1$ and let $u_1, \ldots u_m \in \mathbb{R}$ be such that the inequalities (17), (18) hold for x. Multiply (17) by $f_{\gamma 1}(x)$ and (18) by $f_{\beta 1}(x)$. This implies

$$\begin{aligned} & g_\beta(x)f_{\gamma 1}(x) + \sum_{\mu=2}^{m} f_{\beta\mu}(x)f_{\gamma 1}(x)u_\mu \\ & \quad \le f_{\beta 1}(x)f_{\gamma 1}(x)u_1 \le g_\gamma(x)f_{\beta 1}(x) + \sum_{\mu=2}^{m} f_{\gamma\mu}(x)f_{\beta 1}(x)u_\mu \end{aligned}$$

for $\beta = 1, \ldots, k_1$ and $\gamma = k_1+1, \ldots, k$. Dropping the middle term results in (19) whence $S_1 \subseteq S_2$.

$S_1 \supseteq S_2$:

W.l.o.g. let $S_2 \neq \emptyset$, fix $x \in S_2$ and let $u_2, \ldots u_m \in \mathbb{R}$ be such that the inequalities (19) hold for x. Divide (19) by $f_{\beta 1}(x)$ if $f_{\beta 1}(x) > 0$, and by $f_{\gamma 1}(x)$ if $f_{\gamma 1}(x) > 0$. Unless $f_{\beta 1}(x) = 0$ and $f_{\gamma 1}(x) = 0$ (in this case (19) reads $0 \le 0$ and can be omitted) one gets equivalently

$$g_\beta(x) + \sum_{\mu=2}^{m} f_{\beta\mu}(x)u_\mu \le 0, \qquad \text{if } f_{\beta 1}(x) = 0 \text{ and } f_{\gamma 1}(x) > 0, \tag{20}$$

$$0 \le g_\gamma(x) + \sum_{\mu=2}^{m} f_{\gamma\mu}u_\mu, \qquad \text{if } f_{\beta 1}(x) > 0 \text{ and } f_{\gamma 1}(x) = 0, \tag{21}$$

$$\left(g_\beta(x) + \sum_{\mu=2}^{m} f_{\beta\mu}(x)u_\mu\right) / f_{\beta 1}(x) \le \left(g_\gamma(x) + \sum_{\mu=2}^{m} f_{\gamma\mu}(x)u_\mu\right) / f_{\gamma 1}(x), \quad \text{if } f_{\beta 1}(x) f_{\gamma 1}(x) > 0\,. \tag{22}$$

Due to (16) there exists at least one pair $(\beta^*, \gamma^*) \in \{1, \ldots, k_1\} \times \{k_1+1, \ldots, k\}$ such that $f_{\beta^* 1}(x) f_{\gamma^* 1}(x) > 0$. Let M_1 be the maximum of the left–hand sides of all inequalities (22) and let M_2 be the minimum of the righthand sides of all inequalities (22). Then M_1, M_2 are attained for some indices $\beta = \beta_0$ and $\gamma = \gamma_0$, respectively. Since β and γ vary independently there is an inequality (22) with $\beta = \beta_0$ and $\gamma = \gamma_0$ simultaneously. This proves $M_1 \le M_2$. Choose $u_1 \in [M_1, M_2]$ and apply (20) and (22), respectively, with $\gamma = \gamma_0$ (which implies $f_{\gamma_0 1}(x) > 0$) and $\beta = 1, \ldots, k_1$. If $f_{\beta 1}(x) = 0$ then (20) yields to the corresponding inequality in (17). If $f_{\beta 1}(x) > 0$ then

$$\left(g_\beta(x) + \sum_{\mu=2}^{m} f_{\beta\mu}(x)u_\mu\right) / f_{\beta 1}(x) \le M_1 \le u_1$$

implies the corresponding inequality in (17). By applying (21) and (22), respectively with $\beta = \beta_0$ the inequalities (18) can be seen analogously whence $S_2 \subseteq S_1$.

□

The inequalities in (19) arise by multiplying the corresponding inequalities (17) and (18) by $f_{\gamma 1}(x)$ and $f_{\beta 1}(x)$, respectively. Sometimes it is more convenient to write $f_{\beta 1}(x)$ and $f_{\gamma 1}(x)$ in the form

$$f_{\beta 1}(x) = h_{\beta\gamma}(x)\tilde{f}_{\beta 1}(x), \quad f_{\gamma 1}(x) = h_{\beta\gamma}(x)\tilde{f}_{\gamma 1}(x)$$

with nonnegative functions $\tilde{f}_{\beta 1}$, $\tilde{f}_{\gamma 1}$, $h_{\beta\gamma}$ defined on D. Then the elimination procedure gives some hope that it suffices to multiply (17) and (18) only by $\tilde{f}_{\gamma 1}(x)$ and $\tilde{f}_{\beta 1}(x)$, respectively, in order to end up with the modification

$$g_\beta(x)\tilde{f}_{\gamma 1}(x) + \sum_{\mu=2}^{m} f_{\beta\mu}(x)\tilde{f}_{\gamma 1}(x)u_\mu \le g_\gamma(x)\tilde{f}_{\beta 1}(x) + \sum_{\mu=2}^{m} f_{\gamma\mu}(x)\tilde{f}_{\beta 1}(x)u_\mu \qquad \beta = 1, \ldots, k_1,\ \gamma = k_1+1, \ldots, k, \tag{23}$$

of the corresponding inequality in (19). This multiplication process shows, in particular, that the inclusion

$$S_1 \subseteq S_3 := \{\, x \in D \mid \exists\, u_\mu \in \mathbb{R},\ \mu = 2, \ldots, m \,:\, (23) \text{ holds} \,\}$$

is true. In order to prove $S_3 \subseteq S_1$ fix $x \in S_3$ and choose $u_2, \ldots, u_m \in \mathbb{R}$ such that (23) holds for x. Multiplying the corresponding inequality (23) by $h_{\beta\gamma}(x)$ yields to (19), hence $x \in S_2$, and Theorem 1 implies $x \in S_1$. Thus we have proved the following corollary.

Corollary 1

With the notation and the assumptions of Theorem 1 let

$$f_{\beta 1}(x) = h_{\beta\gamma}(x)\tilde{f}_{\beta 1}(x), \quad f_{\gamma 1}(x) = h_{\beta\gamma}(x)\tilde{f}_{\gamma 1}(x)$$

with nonnegative functions $\tilde{f}_{\beta 1}$, $\tilde{f}_{\gamma 1}$, $h_{\beta\gamma}$ defined on D. Then the assertion of Theorem 1 remains true if $f_{\beta 1}(x)$, $f_{\gamma 1}(x)$ are replaced in (19) by $\tilde{f}_{\beta 1}(x)$ and $\tilde{f}_{\gamma 1}(x)$, respectively.

□

Corollary 1 is particularly useful if $f_{\beta 1} = f_{\gamma 1} \geq 0$ where $f \geq 0$ means $f(x) \geq 0$ for all $x \in D$. Then $h_{\beta\gamma} := f_{\beta 1} = f_{\gamma 1} \geq 0$, $\tilde{f}_{\beta 1} = \tilde{f}_{\gamma 1} := 1 > 0$ and the corresponding inequality in (19) reads

$$g_\beta(x) + \sum_{\mu=2}^{m} f_{\beta\mu}(x)u_\mu \leq g_\gamma(x) + \sum_{\mu=2}^{m} f_{\gamma\mu}(x)u_\mu \,.$$

Another typical application of Corollary 1 occurs if the functions $f_{\lambda\mu}$, g_λ all are polynomials and if $f_{\beta 1}$ and $f_{\gamma 1}$ have a non–constant polynomial as a common factor. We will meet these situations in our subsequent examples.

We remark that no topological assumption such as continuity of $f_{\lambda\mu}$, g_λ or connectivity of D is required in Theorem 1 . The assumption (14) prevents $f_{\lambda 1}$ from being completely omitted in (17), (18) and (19). If $f_{\lambda 1}(x) \leq 0$ on D one can simply fulfill (15) by multiplying the corresponding inequality by -1. If neither $f_{\lambda 1}(x) \geq 0$ nor $f_{\lambda 1}(x) \leq 0$ holds uniformly on D on can split D in several appropriate subdomains D_i with $\bigcup_i D_i = D$ for each of which the assumptions of Theorem 1 hold. The restriction (16) cannot be dropped. This can be seen from the example

$$1 + x_1u_2 \leq x_2u_1, \quad x_1u_1 \leq 1 + x_2u_2, \quad D = O_1 := \{\,(x_1,x_2)\,|\; x_1 \geq 0,\; x_2 \geq 0\,\} \tag{24}$$

where $f_{11}(x) = f_{22}(x) := x_2$, $f_{12}(x) = f_{21}(x) := x_1$, $g_1(x) = g_2(x) := 1$ and where $k = m = n = 2$, $k_1 = 1$. The assumption (16) is not fulfilled for $x = (0,0)$ since $f_{\lambda 1}(0,0) = 0$ for $\lambda \in \{1,2\}$. The inequality (19) reads

$$x_1 + x_1^2u_2 \leq x_2 + x_2^2u_2$$

which is true for $x_1 = x_2 = 0$ while (24) apparently does not hold for $x_1 = x_2 = 0$ and any choice of $u_1 \in \mathbb{R}$.

Note that in our example the functions $f_{\lambda\mu}$ are continuous. Therefore, the equivalence in Theorem 1 apparently cannot be forced by requiring continuity of $f_{\lambda\mu}$, g_λ at the expense of dropping (16). We will illustrate a possible reason in our example. To this end we choose $D := O_1 \backslash \{(0,0)\}$ for the moment. Then (16) holds and Theorem 1 can be applied. Choose $x_1 = x_2 = \varepsilon > 0$. By (24) we get $1 + \varepsilon u_2 = \varepsilon u_1$ whence $u_1 = \dfrac{1}{\varepsilon} + u_2$. Let ε tend to $+0$ which means that (x_1, x_2) approaches the origin in O_1 along the line $x_1 = x_2$. In order to fulfill (24) the two parameters u_1, u_2

must necessarily be chosen in such a way that the absolute value of at least one of them tends to infinity. This situation does, however, not occur in our subsequent considerations since our parameters u_μ will be the matrix entries a_{ij} and the components b_i of the righthand side b of a linear system $Ax = b$. They will be restricted to compact intervals by $A \in [A]$ and $b \in [b]$. This generates inequalities of the form $\underline{a} \le u_\mu \le \overline{a}$ with a corresponding function $f_{\lambda\mu} = 1$. Since such an inequality depends on a single u_μ, it is only used when this parameter is eliminated. Therefore, in the sequel the assumption (16) will be fulfilled for any domain D.

Under appropriate assumptions on the number of the given inequalities and on the signs of the functions $f_{\lambda\mu}$ Theorem 1 and Corollary 1 can be applied successively in order to eliminate some or all expressions in which the parameters u_μ occur linearly. However, the number of inequalities might then increase drastically as already simple examples show.

We shortly summarize the steps to be executed when eliminating the parameters in the inequalities describing some set $S_1 \subseteq D$:

Elimination process

Given a domain $D \subseteq \mathbb{R}^n$ and a set of inequalities in $x \in D$ with parameters $u_1, \ldots u_m$ which occur linearly. Denote D together with this set of inequalities as a record and store it on a stack named Stack 1 .

Step 1

Fetch the first record (i. e., the domain and the corresponding set of inequalities) from Stack 1, fix a parameter, say u_1, bring those inequalities into the form (17), (18) which contain u_1. (Renumber and rename eventually, in order to have a domain named D, a parameter named u_1, and subsequent inequalities according to (17), (18).)

Step 2

Check the assumptions of Theorem 1 for the inequalities which contain u_1. If (15) is not satisfied then multiply the corresponding inequality by -1. If this does not help split D into appropriate subdomains D_i and replace the record with D by corresponding ones with D_i. If (16) is not fulfilled for each D_i then stop. Otherwise put the records to a stack named Stack 2.

Step 3

As long as Stack 2 is not empty fetch from it the last record and eliminate u_1 according to Theorem 1 or Corollary 1. If the new record does no longer contain any parameter u_μ then store it into a file. Otherwise put it to Stack 1 as last element. If Stack 1 is not empty goto Step 1.

$\square$

Now we want to apply Theorem 1 and, whenever possible, Corollary 1 in order to characterize particular subsets of S as announced in Section 1.

3.2 Symmetric linear systems

In order to characterize S_{sym} in (2) we first remark that S_{sym} apparently is empty if $[A] \in \mathbb{R}^{n\times n}$ does not contain a symmetric matrix as an element. If $\underline{A} \neq \underline{A}^T$ or $\overline{A} \neq \overline{A}^T$ we could replace $[A]$ by the largest matrix $[B] \subseteq [A]$ with $[B] = [B]^T$ since $[A]\backslash[B]$ does not contain a symmetric matrix as an element and therefore does not influence S_{sym}. This is the reason why we will assume $[A] = [A]^T$, without loss of generality, from the beginning.

Let O be a fixed orthant. We start with $D = O$ and (4) – (6), this time reducing the amount of free parameters nearly to one half by using $a_{ij} = a_{ji}$. The elimination process for the b_i and the diagonal entries a_{ii} is the same as for S and is left to the reader. The elimination of the off–diagonal entries a_{ij}, $i < j$, $i, j = 1, \dots, n$ differs due to the dependency $a_{ij} = a_{ji}$. For instance, when handling a_{12} first, one gets the (non–trivial) new inequalities

$$\underline{b}_1 - \hat{a}_{11}^+ x_1 - \sum_{j=3}^n a_{1j}x_j \le \hat{a}_{12}^+ x_2, \tag{25}$$

$$\hat{a}_{12}^- x_2 \le \overline{b}_1 - \hat{a}_{11}^- x_1 - \sum_{j=3}^n a_{1j}x_j, \tag{26}$$

$$\underline{b}_2 - \hat{a}_{22}^+ x_2 - \sum_{j=3}^n a_{2j}x_j \le \hat{a}_{12}^+ x_1, \tag{27}$$

$$\hat{a}_{12}^- x_1 \le \overline{b}_2 - \hat{a}_{22}^- x_2 - \sum_{j=3}^n a_{2j}x_j, \tag{28}$$

$$b_1^- x_1 - a_{11}^+ x_1^2 - \sum_{j=3}^n a_{1j}x_1x_j \le b_2^+ x_2 - a_{22}^- x_2^2 - \sum_{j=3}^n a_{2j}x_2x_j, \tag{29}$$

$$b_2^- x_2 - a_{22}^+ x_2^2 - \sum_{j=3}^n a_{2j}x_2x_j \le b_1^+ x_1 - a_{11}^- x_1^2 - \sum_{j=3}^n a_{1j}x_1x_j, \tag{30}$$

where

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

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

$$b_i^- := \begin{cases} \underline{b}_i, & \text{if } x_i \ge 0 \\ \overline{b}_i, & \text{if } x_i < 0 \end{cases}, \qquad b_i^+ := \begin{cases} \overline{b}_i, & \text{if } x_i \ge 0 \\ \underline{b}_i, & \text{if } x_i < 0 \end{cases}.$$

The inequalities (25) – (28) coincide with those for S. The inequalities (29), (30) are new. They contain quadratic polynomials. When eliminating a_{1j} for $j = 3, \dots, n$ according to Corollary 1, the i–th inequality in (4) has to be multiplied by x_i for $i = 3, \dots, n$. Afterwards, no additional multiplication is needed in inequalities which have a form analogous to (29), (30). This is true because the function $f_{\lambda\mu}$ in front of

a_{ij} reads $f_{\lambda\mu}(x) = x_i x_j$ in these inequalities, and in the remaining (non–quadratic) inequalities they are given by $f_{\lambda\mu}(x) = x_i$, $f_{\lambda\mu}(x) = x_j$ and $f_{\lambda\mu}(x) = 1$, respectively. Note that the sign of the function $x_i x_j$ remains constant over a fixed orthant O. This is the reason, why no splitting is needed for $D = O$ during the elimination process. Pursuing this process shows that the final inequalities for $S_{\text{sym}} \cap O$ consist of the inequalities which characterize S, and quadratic inequalities. We thus get the following theorem (see also [1], [2]).

Theorem 2

Let $[A] = [A]^T \in \mathrm{IR}^{n\times n}$ (not necessarily regular) and let $[b] \in \mathrm{IR}^n$. Then in each orthant the symmetric solution set S_{sym} can be represented as the intersection of the unsymmetric solution set S and sets with quadrics as boundaries.

□

Skew–symmetric linear systems and persymmetric linear systems can be handled analogously and yield to a result similar to Theorem 2. For details see [2].

3.3 Hankel and Toeplitz systems

Analogously to Section 3.2 we restrict $\underline{A}$, $\overline{A}$ to be Hankel matrices in order to give some remarks on the solution set

$$S_{\text{Hank}} := \{\, x \in \mathbb{R}^n \,|\, Ax = b,\ A \in [A] \ \text{Hankel matrix},\ b \in [b]\,\} \subseteq S_{\text{sym}} \subseteq S\,.$$

Again we do not require that $[A]$ is regular. This time not only the way but also the results are new and differ essentially from the previous ones. The reason consists in a possible increase of the polynomial degree of $f_{\lambda\mu}$ during the elimination process. In addition, these polynomials have no longer constant sign in a fixed orthant. This can be seen by the following example of a bidiagonal Hankel interval matrix.

Example 1

$$[A] := \begin{pmatrix} 0 & [s] & [d] \\ [s] & [d] & 0 \\ [d] & 0 & \end{pmatrix} \in \mathrm{IR}^{3\times 3}, \qquad [b] \in \mathrm{IR}^3\,.$$

We start with

$$\underline{b}_1 \le s x_2 + d x_3 \le \overline{b}_1, \qquad \underline{b}_2 \le s x_1 + d x_2 \le \overline{b}_2, \qquad \underline{b}_3 \le d x_1 \le \overline{b}_3,$$
$$\underline{d} \le d \le \overline{d}, \qquad \underline{s} \le s \le \overline{s},$$

and, for simplicity, we restrict ourselves to the first orthant O_1, i. e., we apply Corollary 1 with $D = O_1$. After having eliminated the s–terms we obtain

$$\begin{aligned} \underline{b}_1 - d x_3 &\le \overline{s} x_2, & \underline{s} x_2 &\le \overline{b}_1 - d x_3, \\ \underline{b}_2 - d x_2 &\le \overline{s} x_1, & \underline{s} x_1 &\le \overline{b}_2 - d x_2, \\ \underline{b}_1 x_1 - d x_1 x_3 &\le \overline{b}_2 x_2 - d x_2^2, & \underline{b}_2 x_2 - d x_2^2 &\le \overline{b}_1 x_1 - d x_1 x_3, \\ \underline{b}_3 &\le d x_1 \le \overline{b}_3, & \underline{d} &\le d \le \overline{d}, \end{aligned}$$

whence

$$\left\{ \begin{array}{c} \underline{b}_1 - \overline{s}x_2 \le dx_3 \le \overline{b}_1 - \underline{s}x_2, \qquad \underline{b}_2 - \overline{s}x_1 \le dx_2 \le \overline{b}_2 - \underline{s}x_1, \\ \underline{b}_1 x_1 - \overline{b}_2 x_2 \le d(x_1x_3 - x_2^2) \le \overline{b}_1 x_1 - \underline{b}_2 x_2, \\ \underline{b}_3 \le dx_1 \le \overline{b}_3, \qquad \underline{d} \le d \le \overline{d}\,. \end{array} \right\} \tag{31}$$

In order to eliminate the d–terms one has to take into account the signs of the expression $x_1x_3 - x_2^2$. The inequality

$$x_1x_3 - x_2^2 \ge 0. \tag{32}$$

describes a circular cone C which is independent of the coefficients of $[A]$ and $[b]$.[1] Its boundary $x_1x_3 - x_2^2 = 0$ can be rewritten as $x_2^2 + v^2 - u^2 = 0$ where $x_1 = u + v$ and $x_3 = u - v$. The axis of C is given by $(t,0,t)^T$, $t \in \mathbb{R}$; in particular, C lies symmetric with respect to the x_1x_3–plane. It contains the x_1– and the x_3–axis on its surface and divides O_1 into two parts $D_1 := O_1 \cap C$ and $D_2 := \overline{(O_1 \backslash D_1)}$. On D_1 the inequality (32) holds while on D_2 the inequality sign reverses in (32). From (31) we obtain as description of $S_{\text{Hank}} \cap O_1$:

$$\underline{b}_1 - \overline{s}x_2 \le \overline{d}x_3, \quad \underline{b}_2 - \overline{s}x_1 \le \overline{d}x_2, \quad \underline{b}_3 \le \overline{d}x_1,$$

$$\underline{b}_1 x_1 - \overline{s}x_2x_1 \le \overline{b}_3 x_3, \quad \underline{b}_2 x_1 - \overline{s}x_1^2 \le \overline{b}_3 x_2, \quad \underline{b}_1 x_2 - \overline{s}x_2^2 \le \overline{b}_2 x_3 - \underline{s}x_1x_3,$$

$$\begin{array}{c} \underline{b}_1 x_1 - \overline{b}_2 x_2 \le \overline{d}(x_1x_3 - x_2^2), \text{ if } x \in D_1, \\ \underline{b}_1 x_1 - \overline{b}_2 x_2 \le \underline{d}(x_1x_3 - x_2^2), \text{ if } x \in D_2, \end{array}$$

$$\begin{array}{c} \underline{b}_1 x_1^2 - \overline{b}_2 x_1x_2 \le \overline{b}_3(x_1x_3 - x_2^2), \text{ if } x \in D_1, \\ \underline{b}_1 x_1^2 - \overline{b}_2 x_1x_2 \le \underline{b}_3(x_1x_3 - x_2^2), \text{ if } x \in D_2, \end{array}$$

$$\begin{array}{c} \underline{b}_1 x_1x_2 - \overline{b}_2 x_2^2, \le (\overline{b}_2 - \underline{s}x_1)(x_1x_3 - x_2^2), \text{ if } x \in D_1, \\ \underline{b}_1 x_1x_2 - \overline{b}_2 x_2^2 \le (\underline{b}_2 - \overline{s}x_1)(x_1x_3 - x_2^2), \text{ if } x \in D_2, \end{array}$$

$$\begin{array}{c} \underline{b}_1 x_1x_3 - \overline{b}_2 x_2x_3 \le (\overline{b}_1 - \underline{s}x_2)(x_1x_3 - x_2^2), \text{ if } x \in D_1, \\ \underline{b}_1 x_1x_3 - \overline{b}_2 x_2x_3 \le (\underline{b}_1 - \overline{s}x_2)(x_1x_3 - x_2^2), \text{ if } x \in D_2\,. \end{array}$$

We have omitted here the dual inequalities, which are obtained by reversing the inequality signs and by replacing the lower bars by upper ones and vice versa. These inequalities recommend, in particular, that $S_{\text{Hank}} \cap O_1$ should be better replaced by the two subsets $S_{\text{Hank}} \cap D_1$ and $S_{\text{Hank}} \cap D_2$ for each of which the set of inequalities remains fixed. Note that for a complete characterization one has to add the inequalities

$$\begin{array}{rl} x_1x_3 - x_2^2 \ge 0, & \text{(describes } C\text{)} \\ x_i \ge 0 \text{ for } i = 1,\, 2,\, 3 & \text{(describes } O_1\text{)} \end{array}$$

in the case of D_1 and

$$\begin{array}{rl} x_1x_3 - x_2^2 \le 0, & \text{(describes } \overline{\mathbb{R}^3 \backslash C}\text{)} \\ x_i \ge 0 \text{ for } i = 1,\, 2,\, 3 & \text{(describes } O_1\text{)} \end{array}$$

in the case of D_2.

□

[1] Presently we do not know whether this independency always occurs when computing S_{Hank} for more general situations.

We consider now Toeplitz matrices. As can be seen from the definition in Section 2 a Toeplitz matrix A becomes a Hankel matrix if it is multiplied from the left by the permutation matrix E, which has ones in the northeast–southwest diagonal and zeros otherwise. Therefore, the solution set

$$S_{\mathrm{Toep}} := \{\, x \in \mathbb{R}^n \mid Ax = b,\ A \in [A] \text{ Toeplitz matrix},\ b \in [b] \,\} \subseteq S$$

is identical with S_{Hank} formed for EA and Eb. This means that for Toeplitz matrices and for Hankel matrices the same phenomena occur in view of the solution set.

Now we address to the question how complicated can the resulting shape of S_{Toep} be.

When we have a linear interval system with independent coefficients and with a regular interval system matrix, then, due to the Oettli–Prager theorem [7], the solution set is a compact, convex polyhedron in a fixed orthant O (to be more precise, a union of finitely many compact, convex polyhedrons that correspond to different orthants).

In many applications, we are interested only in *some* of the variables $x_1, \ldots, x_n$. In this case, in mathematical terms, we are interested in the *projection* of the solution set on a subspace formed by the desired variables. For interval systems with *independent* coefficients, this projection is a projection of a polyhedron and thus, also a polyhedron.

In [3], we showed that for *arbitrary* interval linear systems with *dependent* coefficients, we can get projections that are described by algebraic dependencies of arbitrarily high degree (we even showed that an arbitrary algebraic set can be thus represented).

A natural question is: if we restrict ourselves to Toeplitz matrices only, how complicated this projection can be? The following simple example shows that for Toeplitz interval matrices we can have, as 2–dimensional projections, curves of degree n at least. To this end let us consider the Toeplitz system $Ax = b$ consisting of the following equations:

$$\begin{aligned} a \cdot x_1 &= 1, \\ -x_1 + a \cdot x_2 &= 1, \\ -x_1 - x_2 + a \cdot x_3 &= 1, \\ &\ \vdots \\ -x_1 - x_2 - \ldots - x_{n-1} + a \cdot x_n &= 1, \end{aligned}$$

where $a \in [1, 2]$. Therefore, a vector $(x_1, \ldots, x_n)^{\mathrm{T}}$ belongs to the solution set if and only if there exists an a for which $a \cdot x_1 = 1$, $-x_1 + a \cdot x_2 = 1$, $-x_1 - x_2 + a \cdot x_3 = 1$, etc.. From these equations, we can explicitly express x_i, $i > 1$, in terms of x_1:

From the first equation, we get $x_1 = 1 / a$; hence, $a = 1 / x_1$.

From the second equation, we get $x_2 = (1 + x_1) / a = (1 + x_1) x_1 = x_1 + x_1^2$; this expression is quadratic in x_1.

Similarly, from the third equation, we get $x_3 = (1 + x_1 + x_2)\,/\,a = (1 + x_1 + x_1 + x_1^2)x_1 = (1 + x_1)^2 x_1 = x_1 + 2x_1^2 + x_1^3$; this expression is cubic in x_1.

...

Finally, for x_n, we get an expression of n–th degree in terms of x_1: $x_n = x_1(1 + x_1)^{n-1} = x_1 + (n-1)x_1^2 + \ldots + x_1^n$.

Thus, when we are only interested in the values of x_1 and x_n, we get a curve of n–th degree.

A similar remark holds for Hankel systems.

It is worth noticing that if we apply Theorem 1 to the interval system above, we get inequalities of degrees less than n. There is no contradiction here, because a set of lower degree can have higher–degree projections: e. g., for a curve described by two second–order equations $x_2 = x_1^2$ and $x_3 = x_2^2$, its projection on (x_1, x_3) has the form $x_3 = x_1^4$ and is, therefore, of fourth order.

In our next example we show that even the unprojected solution set S_{Toep} needs algebraic equations whose order exceeds two.

Example 2

Let

$$[A] := \begin{pmatrix} [d] & 0 & 0 \\ [s] & [d] & 0 \\ [l] & [s] & [d] \end{pmatrix}, \quad [b] := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

with $[d] = [1, 2]$, $[s] = [-4, -2]$ and $[l] = [-8, -4]$. Then each solution of a system $Ax = b$ with a Toeplitz matrix

$$A = \begin{pmatrix} \delta & 0 & 0 \\ \sigma & \delta & 0 \\ \lambda & \sigma & \delta \end{pmatrix} \in [A] \quad \text{and} \quad b = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

is given by

$$x_1 = \frac{1}{\delta} > 0, \tag{33}$$

$$x_2 = -\frac{\sigma}{\delta} x_1 = -\sigma x_1^2 > 0, \tag{34}$$

$$x_3 = -\frac{\sigma}{\delta} x_2 - \frac{\lambda}{\delta} x_1 = \frac{x_2^2}{x_1} - \lambda x_1^2 > 0, \tag{35}$$

with $\delta \in [d]$, $\sigma \in [s]$ and $\lambda \in [l]$. The corresponding set of inequalities reads

$$\frac{1}{2} \le x_1 \le 1, \quad 2x_1^2 \le x_2 \le 4x_1^2, \quad \frac{x_2^2}{x_1} + 4x_1^2 \le x_3 \le \frac{x_2^2}{x_1} + 8x_1^2\,.$$

or, equivalently,

$$\frac{1}{2} \le x_1 \le 1, \quad 2x_1^2 \le x_2 \le 4x_1^2, \quad 4x_1^3 \le x_1x_3 - x_2^2 \le 8x_1^3. \tag{36}$$

Thus S_{Toep} lies completely in O_1; its boundary is part of the two planes $x_1 = \frac{1}{2}$, $x_1 = 4$, part of the two parabolic cylinders $x_2 = 2x_1^2$, $x_2 = 4x_1^2$ and part of the two algebraic surfaces $x_1x_3 - x_2^2 - 4x_1^3 = 0$, $x_1x_3 - x_2^2 - 8x_1^3 = 0$ which are of order three. Notice that (36) was derived by decoupling the parameters in the equalities (33) – (35). Since the last inequality of (36) is the only one which contains x_3 the degree in (36) cannot be reduced. This shows that in the general case the boundary for the solution set of Toeplitz matrices (and therefore also for Hankel matrices) cannot be characterized by means of hyperplanes and quadrics.

□

If $[A]$ is a lower triangular $n \times n$ interval matrix and if $[b]$ is a degenerate interval vector, i.e., $[b] = [b, b]$, then the ideas of Example 2 can be generalized. Using an inductive argument shows that the boundary of the corresponding solution set S_{Toep} is composed by parts of algebraic surfaces which have order n at most, with two of exact order n. A similar remark holds if $[A]$ is an upper triangular matrix, and for S_{Hank} provided that $[A]$ is a triangular matrix with respect to the counterdiagonal. At the moment we do not know how this degree behaves for S_{Toep} and S_{Hank}, respectively, when $[b]$ is non–degenerate or when $[A]$ is not triangular.

3.4 Linear systems with more general dependencies

The elimination process of Section 3.1 can even be applied to systems of linear equations with dependencies according to (3). Such a system (which may be singular) reads

$$g_i(x) + \sum_{\mu=1}^{m} f_{i\mu}(x)u_\mu = 0, \quad i = 1, \dots, n, \tag{37}$$

where

$$\begin{aligned} g_i(x) &:= -b_{i,0} + \sum_{j=1}^{n} a_{ij,0}x_j\,, \\ f_{i\mu}(x) &:= -b_{i,\mu} + \sum_{j=1}^{n} a_{ij,\mu}x_j\,, \\ u_\mu &\in [u]_\mu = [\underline{u}_\mu, \overline{u}_\mu]\,, \\ & i = 1, \dots, n; \quad \mu = 1, \dots, m. \end{aligned} \tag{38}$$

Replace (37) by the equivalent system of inequalities

$$g_i(x) + \sum_{\mu=1}^{m} f_{i\mu}(x)u_\mu \ge 0, \qquad g_i(x) + \sum_{\mu=1}^{m} f_{i\mu}(x)u_\mu \le 0, \quad i = 1, \dots, n, \tag{39}$$

and (38) by

$$\underline{u}_\mu \le u_\mu \le \overline{u}_\mu, \quad \mu = 1, \ldots, m. \tag{40}$$

Then apply the elimination procedure from Section 3.1 to (39), (40) with $D = \mathbb{R}^n$. In this case D is expected to be split into finitely many subdomains D_i in Step 2. Such subdomains certainly exist due to the particular shape of $f_{i\mu}$. (In fact, $f_{i\mu}(x) \ge 0$ determines here a half space in $\mathbb{R}^n$.)

We emphasize that there is an ambiguity in the *order* of eliminating the parameters $u_1, \ldots, u_m$ since we are free to permute the indices. In this respect it is clear by the equivalence of Theorem 1 that for any order and in each stage the inequalities describe the same set, namely the corresponding solution set. But we are not sure whether the inequalities at the end coincide (up to their order of appearence and after having deleted superfluous ones).

References

[1] G. Alefeld, V. Kreinovich and G. Mayer, The Shape of the Symmetric Solution Set, in: R. B. Kearfott and V. Kreinovich, Eds., *Applications of Interval Computations* (Kluwer, Boston, MA, 1996) 61 – 79.

[2] G. Alefeld, V. Kreinovich and G. Mayer, On the Shape of the Symmetric, Persymmetric, and Skew–Symmetric Solution Set, *SIAM J. Matrix Anal. Appl.* **18** (1997) 693 – 705.

[3] G. Alefeld, V. Kreinovich and G. Mayer, The Shape of the Solution Set of Linear Interval Equations with Dependent Coefficients, *Math. Nachr.* **192** (1998) 23 – 26.

[4] G. Alefeld and G. Mayer, The Cholesky Method for Interval Data, *Linear Algebra Appl.* **194** (1993) 161 – 182.

[5] D. J. Hartfiel, Concerning the Solution Set of $Ax = b$ where $P \le A \le Q$ and $p \le b \le q$, *Numer. Math.* **35** (1980) 355 – 359.

[6] C. Jansson, Interval Linear Systems with Symmetric Matrices, Skew–Symmetric Matrices and Dependencies in the Right Hand Side, *Computing* **46** (1991) 265 – 274.

[7] 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) 405 – 409.

[8] J. Rohn, Private Communication, Berlin, 2001.

[9] S. M. Rump, Verification Methods for Dense and Sparse Systems of Equations, in: J. Herzberger, Ed., *Topics in Validated Computations* (Elsevier, Amsterdam, 1994) 63 – 135.

[10] A. Schrijver, *Theory of Linear and Integer Programming* (Wiley, New York, 1986).