Linear and Multilinear Algebra, 1991, Vol. 29(1991), pp. 141–144
Reprints available directly from the publisher
Photocopying permitted by license only

Printed in the United States of America

An Existence Theorem for Systems of Linear Equations

JIRI ROHN
Faculty of Mathematics and Physics, Charles University, Malostranské nám. 25, 118 00 Prague, Czechoslovakia

(*Received August 17, 1989*)

Given is a constructive proof of the following theorem: A system of linear equations has a [nonnegative] solution if and only if each system constructed by replacing each equation by one of the two associated inequalities has a [nonnegative] solution.

Let

$$Ax = b \tag{S}$$

be a system of linear equations with an $m \times n$ matrix A. Denote $Y_m = \{y \in R^m;\ |y_i| = 1 \text{ for each } i\}$, so that Y_m consists of 2^m elements, and for each $y \in Y_m$ let $D_y = \text{diag}\{y_1, \ldots, y_m\}$ (i.e. $(D_y)_{ii} = y_i$ for each i and $(D_y)_{ij} = 0$ for $i \neq j$). Together with (S), we shall consider the family of systems of linear inequalities of the form

$$D_y Ax \leq D_y b \tag{S_y}$$

for all $y \in Y_m$. Obviously, the ith inequality in (S_y) has the form $(Ax)_i \leq b_i$ if $y_i = 1$ and is equivalent to $(Ax)_i \geq b_i$ if $y_i = -1$. It is the purpose of this paper to give a constructive proof of this theorem:

THEOREM 1 *The system (S) has a [nonnegative] solution if and only if (S_y) has a [nonnegative] solution for each $y \in Y_m$.*

The "only if" part is obvious since each solution of (S) also satisfies (S_y) for each $y \in Y_m$. The "if" part is a consequence of the following theorem, which gives a little more:

THEOREM 2 *Let (S_y) have a solution x_y for each $y \in Y_m$. Then (S) has a solution which is a convex combination of the x_y's.*

If all the x_y's are nonnegative, then their convex combination is also a nonnegative vector; this provides for the respective part of Theorem 1. Theorem 2 was proved in [2] in a nonconstructive way using Farkas lemma [1]. We shall show here that a solution to (S) can be constructed from the x_y's algorithmically, although the algorithm itself is not too much efficient, which is no surprise since it must handle 2^m vectors x_y at the outset.

For the description of the algorithm we shall need a special order of elements in Y_m which is defined inductively via the sets Y_j, $j = 1,\ldots,m-1$, in the following way:

(a) the order of Y_1 is -1, 1;
(b) if $y_1,\ldots,y_{2^j}$ is the order of Y_j, then $(y_1,-1),\ldots,(y_{2^j},-1)$, $(y_1,1),\ldots,(y_{2^j},1)$ is the order of Y_{j+1}.

We additionally define $Y_0 = \{1\}$. Further, for any sequence $s_1,\ldots,s_{2h}$ with an even number of elements, each pair s_j, s_{j+h} is called a conjugate pair, $j = 1,\ldots,h$.

We may now formulate the following "cancellation algorithm" for finding a solution to (S) from known solutions x_y to (S_y), $y \in Y_m$:

Algorithm

STEP 0 Form a sequence of vectors $(x_y^{\mathrm{T}},(Ax_y-b)^{\mathrm{T}})^{\mathrm{T}}$ ordered in the order of Y_m.

STEP 1 For each conjugate pair x, x' in the current sequence compute

$$\lambda = \frac{x'_k}{x'_k - x_k} \qquad \text{if} \quad x'_k \neq x_k$$

$$\lambda = 1 \qquad \text{otherwise}$$

where k is the index of the current last entry and set

$$x := \lambda x + (1-\lambda)x'.$$

STEP 2 Cancel the second part of the sequence and in the remaining part delete the last entry of each vector.

STEP 3 If there remains a single vector x, terminate. Otherwise go to Step 1.

Now, both the algorithm and the preceding theorems are justified by this result:

THEOREM 3 *The vector x obtained in Step 3 of the algorithm satisfies $Ax = b$ and $x \in \mathrm{Conv}\{x_y; y \in Y_m\}$.*

Proof The algorithm starts with 2^m vectors of dimension $n+m$ and proceeds by halving the sequence and deleting the last entry, hence it is finite and at the end gives a single n-dimensional vector x. Consider an $(n+j)$-dimensional vector $\tilde{x}$ in a current step of the algorithm before updating (there are 2^j such vectors) and let $y, y \in Y_j$, be a vector which occupies the same position in the order of Y_j as $\tilde{x}$ in the current sequence. Denote $x_y^j = (\tilde{x}_1,\ldots,\tilde{x}_n)^{\mathrm{T}}$ and $r_y^j = (\tilde{x}_{n+1},\ldots,\tilde{x}_{n+j})^{\mathrm{T}}$. We shall prove that for each $j = m,\ldots,1,0$ and each $y \in Y_j$ there holds

$$y_i(Ax_y^j)_i \leq y_i b_i \qquad (i = 1,\ldots,j) \tag{1.1}$$

$$(Ax_y^j)_i = b_i \qquad (i = j+1,\ldots,m) \tag{1.2}$$

$$(r_y^j)_i = (Ax_y^j - b)_i \qquad (i = 1,\ldots,j) \tag{1.3}$$

$$x_y^j \in X, \tag{1.4}$$

where $X = \mathrm{Conv}\{x_y; y \in Y_m\}$. The proof proceeds by induction on $j = m, \ldots, 0$. The case $j = m$ is trivial since $x_y^m = x_y$ for each $y \in Y_m$, hence (1.1) is equivalent to (S_y) and (1.3) follows from the initial construction in Step 0. So assume (1.1)–(1.4) to hold for some $j \in \{1, \ldots, m\}$ and each $y \in Y_j$. Let $y \in Y_{j-1}$. Since, by the order of Y_j, any two conjugate vectors in Y_j differ only in the jth entry, x_y^{j-1} was constructed in Step 1 by

$$x_y^{j-1} = \lambda x_{(y,-1)}^j + (1-\lambda) x_{(y,1)}^j$$

where

$$\lambda = \frac{(r_{(y,1)}^j)_j}{(r_{(y,1)}^j)_j - (r_{(y,-1)}^j)_j} = \frac{(Ax_{(y,1)}^j - b)_j}{(Ax_{(y,1)}^j - b)_j - (Ax_{(y,-1)}^j - b)_j} \in [0,1] \qquad (2)$$

since $(Ax_{(y,1)}^j - b)_j \le 0$ and $(Ax_{(y,-1)}^j - b)_j \ge 0$ due to (1.1). Hence we have

$$y_i(Ax_y^{j-1})_i \le y_i b_i \qquad (i = 1, \ldots, j-1)$$
$$(Ax_y^{j-1})_i = b_i \qquad (i = j+1, \ldots, m)$$

since (1.1) and (1.2), being satisfied by $x_{(y,-1)}^j$ and $x_{(y,1)}^j$, are also satisfied by their convex combination x_y^{j-1}. From (2) we obtain $(Ax_y^{j-1} - b)_j = \lambda(Ax_{(y,-1)}^j - b)_j + (1-\lambda)(Ax_{(y,1)}^j - b)_j = 0$, hence

$$(Ax_y^{j-1})_j = b_j \qquad (3)$$

holds provided the denominator in (2) is nonzero. If $(Ax_{(y,-1)}^j - b)_j = (Ax_{(y,1)}^j - b)_j$, then the common value is both nonnegative and nonpositive, so that $(Ax_{(y,-1)}^j)_j = b_j = (Ax_{(y,1)}^j)_j$ and (3) again holds. From the updating formula in Step 1 we see that $(r_y^{j-1})_i = \lambda(r_{(y,-1)}^j)_i + (1-\lambda)(r_{(y,1)}^j)_i = \lambda(Ax_{(y,-1)}^j - b)_i + (1-\lambda)(Ax_{(y,1)}^j - b)_i = (Ax_y^{j-1} - b)_i$, so that (1.3) also holds for $j-1$. Since $x_{(y,-1)}^j \in X$, $x_{(y,1)}^j \in X$ and X is convex, we get that $x_y^{j-1} \in X$, thus completing the induction.

So for $j = 0$ we obtain from (1.2), (1.4) that $Ax_y^0 = b$, $x_y^0 \in X$ holds for the single remaining n-dimensional vector x_y^0, which is equal to the above x from Step 3. This concludes the proof. ■

To illustrate the algorithm, consider a very simple example:

$$\begin{aligned} x_1 + x_2 - x_3 &= 1 \\ -2x_1 + 3x_2 + x_3 &= 2. \end{aligned} \qquad (4)$$

We may guess the following solutions to the (S_y)'s: $x_{(-1,-1)} = (0,1,0)^{\mathrm{T}}$, $x_{(1,-1)} = (0,0,3)^{\mathrm{T}}$, $x_{(-1,1)} = (2,0,0)^{\mathrm{T}}$, $x_{(1,1)} = (0,0,0)^{\mathrm{T}}$. The performance of the algorithm may be seen from the following scheme, where the arrows indicate the convex combinations of conjugate vectors:

$$\begin{array}{lll}
(0,1,0,0,1)^{\mathrm{T}} & \left(\tfrac{2}{7},\tfrac{6}{7},0,\tfrac{1}{7}\right)^{\mathrm{T}} & \left(\tfrac{3}{11},\tfrac{9}{11},\tfrac{1}{11}\right)^{\mathrm{T}} \\
(0,0,3,-4,1)^{\mathrm{T}} & (0,0,2,-3)^{\mathrm{T}} & \\
(2,0,0,1,-6)^{\mathrm{T}} & & \\
(0,0,0,-1,-2)^{\mathrm{T}} & &
\end{array}$$

The solution to (4) found is $x = (\frac{3}{11}, \frac{9}{11}, \frac{1}{11})^{\mathrm{T}}$.

Although a practical application of the theorems given remains doubtful, they can still be used in some theoretical considerations. Theorem 2 was employed in the proof of the main convex hull theorem in [3] and used for establishing a necessary and sufficient nonsingularity condition for interval matrices in [4].

References

[1] K. G. Murty, *Linear and Combinatorial Programming*, Wiley, New York, 1976.

[2] J. Rohn, Characterization of a linear program in standard form by a family of linear programs with inequality constraints, *Ekon.-mat. obzor* **26** (1990), 71–74.

[3] J. Rohn, Systems of linear interval equations, *Lin. Alg. Appls.* **126** (1989), 39–78.

[4] J. Rohn, Linear interval equations: enclosing and nonsingularity, KAM Series 89/141, Charles University, Prague, 1989.