ЦРНОГОРСКА АКАДЕМИЈА НАУКА И УМЈЕТНОСТИ
ГЛАСНИК ОДЈЕЉЕЊА ПРИРОДНИХ НАУКА, 17, 2007.

ЧЕРНОГОРСКАЯ АКАДЕМИЯ НАУК И ИСКУССТВ
ГЛАСНИК ОТДЕЛЕНИЯ ЕСТЕСТВЕННЫХ НАУК, 17, 2007

THE MONTENEGRIN ACADEMY OF SCIENCES AND ARTS
GLASNIK OF THE SECTION OF NATURAL SCIENCES, 17, 2007.

UDK 519.854

Yu.G.Evtushenko, A.I. Golikov, M. Jaćimović

ON THE THEOREM OF ALTERNATIVE AND APPLICATIONS

A b s t r a c t

We will present Farkas' formulation of the theorem of alternative related to solvability of system of linear inequalities and one review of proofs based on quite different ideas.

O TEOREMI O ALTERNATIVAMA I PRIMJENAMA

I z v o d

U radu ćemo izložiti Farkasovu formulaciju tepreme o alternativama koja se odnosi na rješivost sistema lineranih nejedenačina i pregled dokaza zasnovanih na različitim idejama.

1. INTRODUCTION

The first theorem of alternative related to system of linear equations and inequalities was published in 1873. by P. Gordon. Later, new theorems of alternative were proved, and they were wide used in proving of the existence of solutions in linear algebra and analysis, and in derivation of the necessary conditions of optimality. C.G. Broyden [2] writes that "Theorems of alternatives lie at the heart of the mathematical programming." Interest for these theorems again increased, especially after one paper of G. B. Broyden from 1998. [1], in which he exposed one new proof of these theorems.

In this paper, in section 2, we give four different proofs of the theorem of alternative. The proof based on separation theorem has very simple geometrical interpretation and is probably the most short. However, it can't be considered as a "elementary proof", because in it some "topological arguments" (though simple) are used. In second proof was used Fourier-Motzkin's method of elimination [3] which can be considered as a particular case of the well-known Tarski's theorem on quantifier elimination. This interesting method principally can be used for construction of theorem provers, but the volume of computing grows too fast whit dimension of the problem.

The third proof which belongs to C.G. Broyden [1],is algebraic. Namely, he proved one property of orthogonal matrices from which he derived Farkas's lemma.

In the main part of this paper, in third section, we give a proof based on the duality theory. This proof is an adaption of the one exposed in [4,5]. Finally, we present an application of the theorem on alternative in construction of numerical method for solving system of linear equations.

2. FARKAS'S LEMMA

We will begin with formulation of one statement known as Farkas' lemma.

Theorem 1. (Farkas's lemma) *Let A be a matrix of the order $m \times n$ and b vector-column of the dimension n. Then either*

$$\exists x \geq 0 \textit{ such that } Ax = b \tag{I}$$

$$\exists z \textit{ such that } A^T z \leq 0 \textit{ and } \langle b, z \rangle > 0. \tag{II}$$

G. Farkas was a professor of Theoretical Physics at the University of Kolozsvar (Hungary). He obtained this result while he had been solving the problem of mechanical equilibrium posed by J. Fourier in 1798. He published this results first time in 1898. in Hungarian, but Farkas's best-known exposition of his famous lemma was published only in 1902 in German.

Farkas's lemma can be formulated thus:

$$(A^T x \leq 0 \Rightarrow \langle b, x\rangle \leq 0) \Longleftrightarrow (b \in AR^n_+).$$

If we denote the vectors-columns of matrix A by $a^1, a^2, \ldots, a^m$, we will obtain new equivalent form of this theorem.

The inequality $\langle b, x\rangle \geq 0$ is a consequence of the system of inequalities

$$\langle a_1, x\rangle \geq 0,\ \langle a_1, x\rangle \geq 0, \ldots \langle a_m, x\rangle \geq 0$$

if and only if vector b is a linear combination

$$b = y_1 a_1 + y_2 a_2 + \cdots + y_m a_m$$

with nonnegative coefficients $y_1, y_2, \ldots, y_n$.

Let us remark that as a theorem of alternative we can also consider the following well-known result related to system of linear equations

$$\exists x \text{ such that } Ax = b \tag{Ia}$$

$$\exists z \text{ such that } A^T z = 0 \text{ and } \langle b, z\rangle \neq 0. \tag{IIa}$$

The proofs of Farkas's lemma.

There is a lot of variants of the theorems of alternative and a lot of their proofs. The very well review of the theorems of this kind can be found in [4]. We will present some of the proofs of Farkas's lemma.

We will separately prove that the systems (I) and (II) are not solvable simultaneously. Assume contrary, that there exist $x_0 \in R^n$ and $z_0 \in R^m$ that are some solutions to (I) and (II) respectively. Then, we have

$$0 = \langle x_0, 0\rangle = \langle x_0, A^T z_0\rangle = \langle Ax_0, z_0\rangle = \langle b, z_0\rangle > 0.$$

Thus, we arrive at a contradiction, and the first part of Farkas's lemma is proved.

Now, we will present some proofs of the second part of Farkas's lemma.

Proof 1. Let us propose that the system (I) has no solution. Then $b \notin L := \{Ax : x \geq 0\}$. The set L is convex and closed, so by the separation theorem of closed convex sets, there exists hyperplane $H := \{x : \langle z, x\rangle = \alpha\}$ containing b ($\alpha = \langle z, b\rangle$) such that

$$(\forall y \in L)\ \langle z, y\rangle < \alpha \Rightarrow (\forall x \in R^n_+)\ \langle A^T z, x\rangle < \alpha.$$

This inequality is possible only for $\alpha > 0$ and $A^T z = 0$. So, we conclude that there exists $z \in R^m$, satisfying (II). $\square$

Let us remark that in this very short and elegant geometric proof was omitted the proof that the set $L := \{Ax : x \geq 0\}$ is closed. In addition, separation theorem is intuitive acceptable, but his proof is not so easy. These two facts are the most sensitive part of the geometrical proof of the theorem of alternative.

Note also that the separation theorem can be formulated as a theorem of alternative: *either b is in L or there exists a separating hyperplane.*

Proof 2. In this proof (see [3]) will be used so called Fourier-Motzkin method for variables elimination in linear inequalities. This method can be consider as a particular case of Tarski's quantifier elimination theorem. It can be used for building of a theorem provers for this case. But, even in the case of a system of linear inequalities with only existential quantifiers, the method has very fast grows of the number of computational operation.

Denote by $a_1, a_2, \ldots, a_m$ and $a^1, a^2, \ldots, a^n$ the rows and the columns of the matrix $A = (a_{ij})_{m\times n}$. Then, system (II) can be written in the form

$$\langle a^1, z\rangle \leq 0,\ \langle a^2, z\rangle \leq 0, \ldots, \langle a^n, z\rangle \leq 0, \langle a^{n+1}, z\rangle \leq -\beta < 0,$$

where by a^{n+1} is denoted vector $-b$.

For example, suppose we wish to eliminate the variable z_1 from the above system. Let us denote $I^+ = \{i : a_{1i} > 0\}$, $I^- = \{i : a_{1i} < 0\}$, $I^0 = \{i : a_{1i} = 0\}$. The new system of inequalities will be constructed using the following rules.

For each pair $(k, l) \in I^+ \times I^-$ let us multiply the inequalities $\langle a^k, z\rangle \leq 0$ and $\langle a^l, z\rangle \leq 0$ by $-a_{1l} > 0$ and $a_{1k} > 0$ respectively. Adding these two inequalities, we obtain one new that is not contain the variable z_l. All inequalities obtained on this way will be add to those already in I_0. If I^+ (or I^-) is empty, we simply delete inequalities with indices in I^- (or in I^+). The inequality with indices in I_0 give new system of linear inequalities $Bz' \leq d, z' = (z_2, \ldots, z_n)$. The procedure of elimination of variable z_1 is described.

Let us remark that if $z' = (z'_2, \ldots, z'_n)$ is a solution of the system $Bz' \leq d$, and

$$\max_{l\in I^-} a_{l1}^{-1}(-\sum_{j=2}^{n} a_{lj} + b_l) \leq z_1 \leq \min_{k\in I^+} a_{k1}^{-1}(\sum_{j=2}^{n} a_{kj} - b_k)$$

then $z = (z_1, z') = (z_1, z_2, \ldots, z_n)$ is a solution of the system (II).

Suppose that the system (II): $A^T z \leq 0$, $\langle -b, z\rangle \leq -\beta < 0$ has no solution. Applying Fourier-Motzkin method for the elimination of the variables $z_1, z_2, \ldots, z_n$, one obtains system of inequalities without variables, that is contradictory. This procedure converts the system (II) in inconsistent system

$$\left(\begin{array}{cc} R & q \end{array} \right) \left(\begin{array}{c} A^T \\ -b^T \end{array} \right) z \leq \left(\begin{array}{cc} R & q \end{array} \right) \left(\begin{array}{c} 0 \\ -\beta \end{array} \right),$$

whit $\beta > 0$ and nonnegative elements of the matrix $(R\, q)$. It means that $RA^T - qb^T = 0$, where at least one $q_i \neq 0$. Consequently, there is $x \geq 0$ such that $Ax - b = 0$. $\Box$.

The third proof appeared in G. Broyden's paper [1]. It is based on one property of orthogonal matrices that is referred as a Broyden's theorem.

Broyden's theorem. *If $Q = (q_{ij})_{n\times n}$ is an orthogonal matrix, then there exist a vector $x > 0$ and a unique diagonal matrix $S = diag(s_1, s_2, \ldots, s_n)$ such that $s_i = \pm 1$ and $SQx = x$.*

The Broyden's proof of this theorem is by induction. For $m = 1$ the theorem is trivially true (Q and S are both equal to either $+1$ or -1.) Assume the theorem is true for all orthogonal matrices of order $m \times m$. Let $Q = (q_{ij})_{(m+1)\times(m+1)}$ be an orthogonal matrices and let

$$Q = \left(\begin{array}{cc} P & r \\ q^t & \alpha \end{array} \right),$$

where $P = (p_{ij})_{m\times m}$. If $\alpha = 1$ or $\alpha = -1$, then $r = q = 0$ and the step induction becomes trivial. So, in further we can assume that $|\alpha| < 1$. Since Q is orthogonal matrix,

$$P^T P + q^T q = I,\ P^T r + \alpha q = 0,\ r^T r + \alpha^2 = 1.$$

From these equations it follows that the matrices

$$Q_1 = P - \frac{rq^T}{\alpha - 1},\ Q_1 = P - \frac{rq^T}{\alpha + 1}$$

are orthogonal and that

$$Q_2^T Q_1 = I - q\frac{2}{1-\alpha^2}q^T.$$

Using induction assumption we conclude that there are $x_1 > 0$ and $x_2 > 0$ and diagonal sign matrices S_1 and S_2 such that $S_1 Q_1 x_1 = x_1$, $S_2 Q_2 x_2 = x_2$. From this we obtain

$$\langle S_1 x_1, S_2 x_2\rangle = \langle Q_1 x_1, Q_2 x_2\rangle = \langle Q_2^T Q_1 x_1, x_2\rangle = \langle x_1, x_2\rangle - \frac{2}{1-\alpha^2}\langle x_1, q\rangle\langle x_2, q\rangle.$$

From this moment we have to consider two cases.

Case 1. If $S_1 \neq S_2$, then $\langle S_1x_1, S_2x_2\rangle < \langle x_1, x_2\rangle$. So, $\langle q, x_1\rangle \neq 0$ and $\langle q, x_2\rangle \neq 0$ and both scalar products have the some signs. For

$$\eta_1 = \frac{-\langle q, x_1\rangle}{\alpha - 1},\ \eta_2 = \frac{-\langle q, x_2\rangle}{\alpha + 1},\ z_1 = \begin{pmatrix} x_1 \\ \eta_1 \end{pmatrix},\ z_2 = \begin{pmatrix} x_2 \\ -\eta_2 \end{pmatrix},$$

$$S^1 = \begin{pmatrix} S_1 & 0 \\ 0 & 1 \end{pmatrix},\ S^2 = \begin{pmatrix} S_2 & 0 \\ 0 & -1 \end{pmatrix},$$

we have

$$Qz_1 = S^1 z_1,\ Qz_2 = S^2 z_2.$$

Now, since $|\alpha| < 1$, and both scalar products $\langle q, x_i\rangle$ have the some signs, one of η_i is positive and one of the vectors z_1 and z_2 is the required vector. In case 1 the proof is completed.

Case 2. If $S_1 = S_2$, then $\langle S_1x_1, S_2x_2\rangle = \langle x_1, x_2\rangle$ and at least one of $\langle q, x_1\rangle$ and $\langle q, x_2\rangle$ is zero. We will assume that $\langle q, x_1\rangle \neq 0$ and $\langle q, x_2\rangle = 0$. So $Px_1 = S_1x_1 = Q_1x_1$ and $Qz_1 = S^1z_1$, where

$$z_1 = \begin{pmatrix} x_1 \\ 0 \end{pmatrix},\ S^1 = \begin{pmatrix} S_1 & 0 \\ 0 & \sigma \end{pmatrix},\ \sigma = \pm 1 \text{ may be chosen arbitrarily.}$$

Now, if we rewrite Q in the form

$$Q = \begin{pmatrix} \alpha_1 & q_1^T \\ r_1 & P_1 \end{pmatrix}$$

where P_1 is a matrix of the type $m \times m$, and repeat the previous argument, we will obtain that there exist a positive vector x_2 and diagonal matrix S_2 with ± 1 on diagonal, such that

$$Qz_2 = S^2 z_2,\ z_2 = \begin{pmatrix} 0 \\ x_2 \end{pmatrix},\ S^2 = \begin{pmatrix} 0 & \sigma' \\ S_1 & 0 \end{pmatrix},$$

where $\sigma' = \pm 1$, $\sigma = \pm 1$ may be chosen arbitrarily.

Combining the equation $Qz_1 = S^1z_1$ and $Qz_2 = S^2z_2$, we obtain

$$Q(z_1 + z_2) = S^1 z_1 + S^2 z_2,$$

with strictly positive coordinate z_{1_j} and z_{2_j} for $j \geq 2$. If for some $j \in \{2, \dots, m\}$ the corresponding diagonal elements of S^1 and S^2 are different, then $\|S^1z_1 + S^2z_2\| < \|z_1 + z_2\| = \|Q(z_1 + z_2)\|$, but it is a contradiction with the previous equality. We can choose the elements σ and σ' so that

$S^1 = S^2$. So, we have $Q(z_1 + z_2) = S^1(z_1 + z_2)$, and since $z_1 + z_2$ is strictly positive and S^1 is a diagonal matrix with ± 1 on diagonal, this establish the induction in case $S_1 = S_2$.

Assume that there exist two positive vectors x and y and two diagonal matrices S and R with ± 1 on diagonal, such that $Qx = Sx = Ry$, and $S \neq R$. Then

$$\langle x, y\rangle = \langle Qx, Qy\rangle = \langle Sx, Ry\rangle < \langle x, y\rangle,$$

giving contradiction. Therefore, $R = S$. This completes the proof of Broyden's theorem. □

The next result that is known as Tucker's theorem is a simple consequence of Broyden's theorem.

Tucker's theorem.*Let A be an skew-symmetric matrix. Then there exist $y \geq 0$ such that $Ay \geq 0$ and $y + Ay > 0$.*

Proof of Tucker's theorem. Since A is skew-symmetric then $(I+A)^{-1}(I-A)$ is orthogonal, so that there exist a positive vector x and unique matrix S such that

$$(I - A)^{-1}(I + A)x = Sx \Leftrightarrow x + Ax = Sx - ASx.$$

If we denote $y = x+Sx$, $z = Ay = Ax+ASx = x-Sx$, then every coordinate y_j of the vector y is equal either $2x_j$ or zero, so $y \geq 0$. Similarly, $z \geq 0$. But $y + z = y + Ay = 2x > 0$. □

Proof 3. Broyden's proof of Farkas's lema Apply Tucker's theorem to skew-symmetric matrix

$$B = \begin{pmatrix} O & O & A & -b \\ O & O & -A & b \\ -A^T & A^T & O & 0 \\ b^T & -b^T & O^T & 0 \end{pmatrix}$$

By Tucker's theorem, there exists a positive vector $y = (z_1, z_2, x, t)^T$ such that $y + By > 0$. Consider the two cases: $t > 0$ and $t = 0$ with $z = z_1 - z_2$. If $t > 0$ the vector y may be normalized so that $t = 1$, from which we obtain $Ax = b$. If If $t = 0$ then $A^T z \leq 0$ and $\langle b, z\rangle > 0$.

In his paper Broyden discussed the question of the relation between Tucker's theorem, Farkas's lemma and Broyden's theorem. He derived Tucker's theorem and Farkas's lemma as a simple consequence of Broyden's theorem. In [8] Ross and Terlaky shown that the Broyden theorem is also a simple consequence of Tucker's theorem and Farkas's lemma. It means that Farkas's lemma, Tucker's theorem and Broyden's theorem are equivalent results.

In [8] was made the observation that for given orthogonal matrix Q the existence of an unique diagonal matrix S with diagonal elements $s_{ii} = \pm 1$ and positive vectors x, such that $Qx = Sx$ is equivalent to existence of a positive vector x such that

$$|Qx| = x,\ x > 0 \Leftrightarrow -x \leq Qx \leq x,\ x > 0,$$

where $|y|$ denotes the vector of the absolute values of the entries of vector y. So, if one finds a vector x satisfying the previous conditions, then $s_{ii} = 1$ if $x_i > 0$ and $s_{ii} = -1$ if $x_i < 0$.

This Roos's and Terlaky's comment was an very perspicuous answer on Broyden's remark [1] that "However it may yet be possible to construct such (for determining of sign matrix) algorithm and the author suspects that if this is the case then any successful example would have more than a passing resemblance to the interior point algorithms, but only the passage of time will resolve this conjecture".

3. DUALITY THEORY AND THEOREMS OF THE ALTERNATIVE

In this part, we follows the ideas presented in [4] and [5].

Let us consider the system

$$Ax = b,\ x \geq 0. \tag{I}$$

Denote the set of solution of this system by X:

$$X = \{x \in R^n : Ax = b, x \geq 0\}.$$

Note that it can be viewed as a feasible set in linear programming problem of minimization of $\langle c, x\rangle$ with $c = 0$. Then the dual problem is the maximization of $\langle b, z\rangle$ on the set Z defined by

$$Z = \{z \in R^m : A^T z \leq 0\} \tag{I'}$$

We will say that it is *adjoint system* to system $Ax = b, x \geq 0$.

Consider also the system

$$A^T u \leq 0, \langle b, u\rangle = \rho, \tag{II}$$

where $\rho > 0$ is arbitrary fixed positive number; the set of solutions of (II) will be denoted by U :

$$U = \{u : A^T u \leq 0, \langle b, u\rangle = \rho\}.$$

It is easy to see that $U \subseteq Z$.

The adjoint system to (II) is

$$Aw - tb = 0, w \geq 0, t \in R, \qquad (II').$$

The sets

$$Z = \{z \in R^m : A^T z \leq 0\} \text{ and } W = \{(w,t) \in R^n \, x \, R : Aw - tb = 0, w \geq 0\}$$

of the solutions to the system (I') and (II') are nonempty, because they contain trivial solutions. The set W can be written as

$$W = \cup_{t \in R}(W_t, t), \text{ where } W_t = \{w \in R^n_+ : Aw = tb\}.$$

It is easy to see that if the system (II) is solvable for a certain $\rho = \rho_0 > 0$, then it is solvable for all $\rho > 0$.

Let us defined the penalties

$$\text{pen}(x, X) = \frac{1}{2}\|b - Ax\|^2, \, x \geq 0, \text{ pen}(u, U) = \frac{1}{2}\|(A^T u)_+\|^2 + \frac{1}{2}(\rho - \langle b, u \rangle)^2,$$

for the violation of the conditions $x \in X$ and $u \in U$. Here, v_+ is a vector whose $i - th$ coordinate is equal to that of the vector v if the later is nonnegative and is zero otherwise. Consider the problems

$$\min\{\text{pen}(x, X) : x \in R^n\} = \min\{\frac{1}{2}\|b - Ax\|^2 : x \geq 0\} \qquad (*)$$

$$\min\{\text{pen}(u, U) : u \in R^m\} = \min\{\frac{1}{2}\|(A^T u)_+\|^2 + \frac{1}{2}(\rho - \langle b, u \rangle)^2 : u \in R^m\} \qquad (**)$$

The problems (I) and (II) may be inconsistent. It means that the sets X and U can be empty. But, the following problems

$$\max\{\langle b, z \rangle - \frac{1}{2}\|z\|^2 : z \in Z\} = \max\{\frac{1}{2}\|b\|^2 - \frac{1}{2}\|z - b\|^2 : z \in Z\}, \quad (*')$$

$$\max\{\rho t - \frac{1}{2}\|w\|^2 - \frac{1}{2}t^2 : (t, w) \in W\} = \max\{\frac{1}{2}\rho^2 - \frac{1}{2}|\rho - t|^2 - \frac{1}{2}\|w\|^2 : (t, w) \in W\} \quad (**')$$

always have unique solutions. Let us remark that these problems are connected with the projection on the sets Z and W : problem (*') is a projecting of the point b on the set Z, while (**') is a projecting of the point $(0, \rho)$ on the set W.

We will prove that pairs (*), (*'), and (**), (**') can be considered as a pairs of primal-dual problems.

It is easy to see that the dual problem of (*') is (*). We shall show that (*) can be viewed as a dual of (*'). Let us introduce a vector $y = b - Ax$. Then (1) can be written in the form

$$\min\{\frac{1}{2}\|y\|^2 : Ax + y = b, y \in R^m, x \geq 0\}. \qquad (*1)$$

Lagrange's function for this problem is given by

$$L(x, y, z) = \frac{1}{2}\|y\|^2 + \langle z, b - Ax - y\rangle = \frac{1}{2}\|y\|^2 + \langle z, b - y\rangle - \langle A^T z, x\rangle, x \geq 0.$$

Then the dual problem is a problem of maximization of the function

$$F(z) = \min\{L(x, y, z) : y \in R^n, x \geq 0\}.$$

The optimality conditions for this problem are:

$$A^T z \leq 0,\ y = -z,\ D(x)A^T Z = 0.$$

So, we obtain that for $z \in Z$

$$F(z) = \frac{1}{2}\|z\|^2 + \langle b, z\rangle - \|z\|^2 =\rangle - \frac{1}{2}\|z\|^2,$$

and the dual of (*) is (*').

Now we can prove the following set of equalities.

Lemma 1. Any solution x^* to problem (*) determines a unique solution z^* to problem (') as $z^* = b - Ax^*$, and it holds that*

$$\|z^*\|^2 = \langle b, z\rangle,\ \langle z^*, Ax^*\rangle = 0,$$

$$\|z^*\| = \mathrm{dist}(x^*, X),\ \|b - z^*\| = \mathrm{dist}(b, Z), \|z^*\|^2 + \|b - z^*\|^2 = \|b\|^2.$$

If x^* is a solution to (*), and $z^* = b - Ax^*$, then $\|z^*\| = \|b - Ax^*\| = \mathrm{dist}(b, X)$ and $\| - z^*\| = \mathrm{dist}(b, Z)$. In addition, the optimality conditions for (*) at the point x^* are

$$-A^T(b - Ax^*) \geq 0,\ x^* \geq 0,\ D(x^*)A^T(b - Ax^*) = 0.$$

and they can be written as

$$A^T z^* \leq 0,\ D(x^*)A^T z^* = 0, x^* \geq 0.$$

It follows that $z^* \in Z$ and

$\langle z^*, Ax^*\rangle = \langle A^T z^*, x^*\rangle = 0$ and $\langle b, z^*\rangle = \langle z^*+Ax^*, z^*\rangle = \|z^*\|^2+\langle x^*, A^T z^*\rangle = \|z^*\|^2$.

$$\langle z^*, Ax^*\rangle = \langle A^T z^*, x^*\rangle = 0,$$
$$\|z^*\|^2 = \langle z^*, b - Ax^*\rangle = \langle b, z^*\rangle - \langle A^T z^*, x^*\rangle = \langle b, z^*\rangle,$$
$$\langle z^*, b - z^*\rangle = 0.$$

As a consequence of the equality $\langle b, z^*\rangle = \|z^*\|^2$, we have

$$\|z^*\| = \|b - Ax^*\| = \mathrm{dist}(x^*, X),\ \|b - z_*\| = \mathrm{dist}(b, Z)$$

and

$$\|z^*\|^2 + \|b - z^*\|^2 = \langle b, z^*\rangle + \|b\|^2 + \langle b, z^*\rangle - 2\langle b, z^*\rangle = \|b\|^2.$$

□

The vector z^* is called *the minimal residual vector* while the vector x^* is a *pseudosolution* to system $Ax = 0, x \geq 0$.

The relation between (**) and (**') are similar as the relation between (*) and (*'), but somewhat different.

Lemma 2. *Any solution u^* to problem (**) determines a unique solution (w^*, t^*) to problem (**') and it holds*

$$w^* = (A^T u^*)_+,\ t^* = \rho - \langle b, u^*\rangle, -\langle w^*, A^T u^*\rangle + t^*\langle b, u^*\rangle = 0,$$

$$\|(A^T u^*)_+\|^2 + (\rho - \langle b, u\rangle)^2 + \mathrm{dist}^2((0, \rho), W) = \rho^2.$$

Proof. The problem (**') is the problem of projection of the vector $(0, \rho)$ onto the set $W = \{(w, t) : Aw - tb = 0, w \geq 0\} \neq \emptyset$ and $(w^*, t^*) = Pr((0, \rho), W)$. For this problem, which has a unique solution, Lagrange function is as follows

$$L(w, t, u) = \rho t - \frac{1}{2}(\|w\|^2 + t^2) - \langle u, tb - Aw\rangle,\ w \geq 0.$$

The optimality conditions (Kuhn-Tucker conditions) at saddle point (w^*, u^*) are

$$-w^* + A^T u^* \leq 0,\ w^* \geq 0,\ D(w^*)(-w^* + A^T u^*) = 0,$$

$$\rho - t^* - \langle b, u^*\rangle = 0, Aw^* - t^*b = 0.$$

It follows from here that $(w^*, t^*) \in W$ and (w^*, t^*) can be written in terms of u^* by

$$w^* = (A^T u^*)_+,\ t^* = \rho - \langle b, u^* \rangle.$$

Using these equalities we obtain

$$A(A^T u^*)_+ - (\rho - \langle b, u^* \rangle) b = 0.$$

This is the optimality condition for problem (*) at the point u^*. So, we can conclude that the vectors (w^*, t^*) and u^* solve the problems (**') and (**), respectively.

Further, using the equality of the values of the primal and dual problems (**) and (**'), and $w^* = (A^T u^*)_+, t^* = \rho - \langle b, u^* \rangle$, one obtains $\|w^*\|^2 + (t^*)^2 = \rho t_*$. Now, the equality $-\langle w^*, A^T u^* \rangle + t^* \langle b, u^* \rangle = 0$ is a direct consequence of the optimality conditions $w^* \geq 0$, $D(w^*)(-w^* + A^T u^*) = 0, \rho - t^* - \langle b, u^* \rangle = 0, Aw^* - t^* b = 0$. Finally, we have

$$\|(A^T u^*)_+\|^2 + (\rho - \langle b, u^* \rangle)^2 = 2\rho t^* - \|w^*\|^2 - (t^*)^2.$$

Taking into account

$$\|w^*\| + (t^*)^2 = \|(A^T u^*)_+\|^2 + (\rho - \langle b, u^* \rangle)^2,\ \|w^*\|^2 + (t^* - \rho)^2 = \text{dist}^2((0, \rho), W)$$

we obtain

$$\|(A^T u^*)_+\|^2 + (\rho - \langle b, u \rangle)^2 + \text{dist}^2((0, \rho), W) = \rho^2.$$

□

Let us remark that as a consequence of $\|w^*\|^2 + (t^*)^2 + \text{dist}^2(0, \rho), W) = \rho^2$ we obtain the estimation $\|w^*\|^2 + (t^*)^2 \leq \rho^2$ and $0 \leq t^* \leq \rho$. The equality $(t^*)^2 + \|w\|^2 = \rho t^*$ can be written as the quadratic equation $(t^*)^2 - \rho t^* + \|w\|^2 = 0$, that in case $w \neq 0$, has real positive solution. This means that the discriminant $D = \rho^2 - 4\|w^*\|^2 \geq 0$, and $\|w^*\|^2 \leq \frac{\rho^2}{4}$.

Farkas's lemma is contained in the following theorem.

Theorem 2. *Let x^* and u^* be arbitrary solutions to problems (*) and (**), respectively, and let the minimum residual vectors z^* and w^* be defined by $z^* = b - Ax^*$ and $w^* = (A^T u^*)_+$. Then the following assertions are valid:*

(i) Only one of of the system (I) and (II) is solvable.

(ii) If system (I) is inconsistent, then the normal solution $\tilde{u}^*$ to system (II) and the minimum residual vector z^* of system (I) are collinear and

$$\tilde{u}^* = \frac{\rho z^*}{\|z^*\|^2},\ z^* = \frac{\rho \tilde{u}^*}{\|\tilde{u}^*\|^2};$$

(iii) If system (II) is inconsistent, then the normal solution $\tilde{x}^*$ to system (I) is collinear to w^* and

$$\tilde{x}^* = \frac{w^*}{t^*}.$$

Proof. Let us recall that the systems (I) and (II) cannot be consistent simultaneously. Let us show that one of them must be consistent. Consider two possibilities separately.

If $X = \emptyset$, then $\text{pen}(x^*, X) > 0$, and $\|z^*\| \neq 0$. Let us consider vector $\tilde{u}^* = \frac{\rho z^*}{\|z^*\|^2}$. Multiplying both sides of this equation by b, and taking into account $\|z^*\|^2 = \langle b, z^* \rangle$, we obtain $\langle b, \tilde{u}^* \rangle = \rho$. In addition, $A^T \tilde{u}^* \leq 0$; hence $U \neq \emptyset$. We have to prove that $\tilde{u}^*$ is the normal solution to (II),i.e. that $\tilde{u}*$ is a solution to the problem

$$\min \left\{ \frac{\|u\|^2}{2 : u \in U} \right\}, \; U = \{u \in R^m : A^T u \leq 0, \langle b, u \rangle = \rho\}, \; \rho > 0.$$

The Lagrange's function for this problem can be written as

$$L(u, \hat{x}) = \frac{\|u\|^2}{2} + \langle \hat{x}, A^T u \rangle + \lambda(\rho - \langle b, u \rangle), \; \tilde{x} \in R^n_+, \; \lambda \in R,$$

and the dual problem is

$$\max\{\rho\lambda - \|b\lambda - A\hat{x}\|^2 : x \geq 0, \; \lambda \in \mathbb{R}\}.$$

Let $(u^*, \hat{x}^*, \lambda^*)$ be a saddle point of this Lagrange's function. Then u^* is a solution to the primal and (x^*, λ^*) is a solution to the dual problem. So, the Khun-Tucker conditions at saddle point $(u^*, \hat{x}^*, \lambda^*)$, are

$$u^* = -A\hat{x}^* + \lambda^* b, \; A^T u^* \leq 0, \; \hat{x}^* \geq 0, \; D(\hat{x}^*), A^T u^* \rangle = 0, \; \rho - \langle b, u^* \rangle = 0.$$

The optimal values of the primal and dual problems ($\frac{1}{2}\|u^*\|^2$ and $\rho\lambda^* - \frac{1}{2}\|u^*\|^2$) are equal, so $\|u^*\|^2 = \rho\lambda^*$. Since $U \neq \emptyset$, $u^* \in U$ and $\langle b, u^* \rangle = \rho > 0$, it holds $u^* \neq 0$ and $\lambda^* > 0$. As a consequence of the previous equalities, we obtain

$$u^* = \lambda^* z^*, \; \hat{x}^* = \lambda^* x^*, \; z^* = b - Ax^*$$

and it is easy to see that the pair (z^*, x^*) satisfies Khun-Tucker conditions for (*'). As a consequence of this relations we obtain

$$u^* = \lambda^* z^* = \frac{\rho z^*}{\|z^*\|^2} = \tilde{u}^*, \; \frac{\rho}{\lambda^*} - \langle b, z^* \rangle = \frac{\rho}{\lambda^*} - \|z^*\|^2.$$

Hence, for $\lambda^* = \frac{\rho}{\|z^*\|^2}$, vector $u^* = \lambda^* z^* = \tilde{u^*}$ is normal solution to (II)

Now, it is easy to prove that $\rho = \|\tilde{u}^*\|\|z^*\|$ and $z^* = \frac{\rho\tilde{u}^*}{\|\tilde{u}^*\|^2}$.

Let us consider the case $U = \emptyset$. Then $\mathrm{pen}(u^*, U) \neq 0$ and $w^* = (A^T u^*)_+ \neq 0$. It follows that $x^* = \frac{w^*}{t^*}$ is a solution to (I). Let $\tilde{x}^*$ be the normal solution to (I),i.e. a solution to the problem

$$\min\left\{\frac{\|x\|^2}{2} : x \in X\right\}$$

The Lagrange's function for this problem is

$$L(x, \mu) = \frac{1}{2}\|x\|^2 + \langle \mu, b - Ax\rangle,$$

and the Khun-Tucker conditions at the saddle point $(\tilde{x}^*, \mu^*)$ are

$$\tilde{x}^* - A^*\mu^* \geq 0,\ \tilde{x}^* \geq 0,\ D(\tilde{x}^*)(\tilde{x}^* - A^T\mu^*) = 0, b - A\tilde{x}^* = 0.$$

But, from here, with $x^* = \frac{w^*}{t^*}$, $\mu^* = \frac{u^*}{t^*}$ we obtain the Khun-Tucker optimality conditions for problem (**'). Hence, $\tilde{x}^* = \frac{w^*}{t^*}$ is the normal solution to system (I). □

Thus, this Theorem reduces problem of solvability of systems (I) (system (II)) to minimizing of the residual of system (II) (system (I)). If minimal residual vector of one system is nonzero, then this system is inconsistent and the residual can be used in simple formulas to find the normal solution to the corresponding consistent system.

REFERENCES

[1] C.G. Broyden, A simple algebaric proof of Farkas's lemma and related theorems, Optimization methods and software, 1998., vol. 8(3-4): pp. 185-199.

[2] C.G. Broyden, On theorems of the aternative, Optimization methods and software, 2001., vol 16, pp. 101-111

[3] V.Chandru, C. Lassez, J.L. Lassez, Qualitative theorem proving in linear constraints, in book: Verification: Theory and Practice, Lecture notes in computer science, Springer, Berlin/Neideleberg, 2004, pp. 395-406

[4]. A. Dax, V.P. Sreedharan, Theorems of the alternative and duality, JOTA, 1994, No.3, pp. 561-590.

[5]. Yu.G. Evtushenko,A.I.Golikov, Theorems of alternative and their applications in numerical methods, Computational Mathematics and Mathematical Physics, 2003, vol. 43., No.3, pp. 338-358.

[6]. A.I.Golikov, Yu.G. Evtushenko, New perspective of the theorem of alternatives, High performance algorithm and software for nonlinear optimization, Kluwer Academic publisher B.V, 2004, pp. 227-241.

[7]. A.I.Golikov, Yu.G. Evtushenko, M.Jaćimović: Normal solution and regularization of linear programming problem, Dynamic of nonhomogeneous system, (ed. Yu.S. Popkov), Editorial URSS, 2004., No.8, (in Russian)

[8]. C. Roos, T. Terlaky, Note on a paper of Broyden, Operations Research Letters, 1999, vol. 25, no. 4, pp. 183-186.