Doklady Mathematics, Vol. 64, No. 3, 2001, pp. 370–373. Translated from Doklady Akademii Nauk, Vol. 381, No. 4, 2001, pp. 444–447.

A New Method for Solving Systems of Linear Equalities and Inequalities

A. I. Golikov and Corresponding Member of the RAS Yu. G. Evtushenko

Received July 26, 2001

Solving systems of equalities and inequalities has an extensive literature. We only refer to [1–3]. Usually, these problems are solved by reduction to the unconstrained minimization of the residuals of the original system. Every linear system is associated with an alternative system such that one and only one of these systems is solvable. It is not known a priori whether a given system has solutions. Thus, the problem is, first, to determine whether the given system is solvable and, second, to find its solution if it is solvable.

In this paper, we suggest a method for solving linear systems based on alternative theorems [3–6]. Given a linear system, we construct an alternative system such that the dimension of its variables equals the number of equalities and inequalities (except constraints on the signs of variables) in the original system. The method for solving the original solvable system consists in minimizing the residual of the alternative inconsistent system. According to the result of this minimization, we determine a normal solution (a solution with minimum Euclidean norm) to the original system. The substantiation of the method is based on the duality theory.

Since the dimensions of the variables in the considered systems are different, the passage from the original consistent system to the minimization problem for the residual of the alternative inconsistent system may be very expedient. This reduction may lead to the problem of minimization with respect to variables of lower dimensions and make it possible to determine a normal solution to the original system.

Let A be an $m \times n$ matrix given in the form

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

Here, A_{11}, A_{12}, A_{21}, and A_{22} are rectangular matrices of sizes $m_1 \times n_1$, $m_1 \times n_2$, $m_2 \times n_1$, and $m_2 \times n_2$, respectively. Suppose that vectors $x \in R^n$, $u \in R^m$, and $b \in R^m$ admit decompositions $x^{\mathrm{T}} = [x_1^{\mathrm{T}}, x_2^{\mathrm{T}}]$, $u^{\mathrm{T}} = [u_1^{\mathrm{T}}, u_2^{\mathrm{T}}]$, and $b^{\mathrm{T}} = [b_1^{\mathrm{T}}, b_2^{\mathrm{T}}]$, where $x_1 \in R^{n_1}$, $x_2 \in R^{n_2}$, $n = n_1 + n_2$, $u_1 \in R^{m_1}$, $u_2 \in R^{m_2}$, $b_1 \in R^{m_1}$, $b_2 \in R^{m_2}$, and $m = m_1 + m_2$. Let us introduce the auxiliary sets

$$\Pi_x = \{[x_1, x_2]: x_1 \in R_+^{n_1}, x_2 \in R^{n_2}\},$$

$$\Pi_u = \{[u_1, u_2]: u_1 \in R_+^{m_1}, u_2 \in R^{m_2}\}.$$

Consider the system of linear equalities and inequalities

$$\begin{aligned} A_{11}x_1 + A_{12}x_2 &\ge b_1, \\ A_{21}x_1 + A_{22}x_2 &= b_2, \\ x_1 &\ge 0_{n_1}. \end{aligned} \tag{I}$$

We define the system conjugate to (I) as

$$\begin{aligned} A_{11}^{\mathrm{T}}z_1 + A_{21}^{\mathrm{T}}z_2 &\le 0_{n_1}, \\ A_{12}^{\mathrm{T}}z_1 + A_{22}^{\mathrm{T}}z_2 &= 0_{n_2}, \\ z_1 &\ge 0_{m_1}. \end{aligned} \tag{I'}$$

and the system alternative to (I) as

$$\begin{aligned} A_{11}^{\mathrm{T}}u_1 + A_{21}^{\mathrm{T}}u_2 &\le 0_{n_1}, \\ A_{12}^{\mathrm{T}}u_1 + A_{22}^{\mathrm{T}}u_2 &= 0_{n_2}, \\ b_1^{\mathrm{T}}u_1 + b_2^{\mathrm{T}}u_2 &= \rho, \\ u_1 &\ge 0_{m_1}. \end{aligned} \tag{II}$$

Here, $\rho > 0$ is an arbitrary fixed positive number. Note that the requirement that ρ be positive automatically implies that $\|b\| \ne 0$.

Let us introduce a vector $w \in R^{n+1}$ representable in the form

$$w^{\mathrm{T}} = [w_1^{\mathrm{T}}, w_2^{\mathrm{T}}, w_3],$$

where $w_1 \in R^{n_1}$, $w_2 \in R^{n_2}$, and $w_3 \in R^1$, and the auxiliary set

$$\Pi_w = \{[w_1, w_2, w_3]: w_1 \in R_+^{n_1}, w_2 \in R^{n_2}, w_3 \in R^1\}.$$

Computing Center, Russian Academy of Sciences, ul. Vavilova 40, GSP-1, Moscow, 117967 Russia
e-mail: evt@ccas.ru, gol@ccas.ru

The system conjugate to (II) has the form

$$\begin{aligned} A_{11}w_1 + A_{12}w_2 - b_1w_3 &\geq 0_{m_1}, \\ A_{21}w_1 + A_{22}w_2 - b_2w_3 &= 0_{m_2}, \\ w_1 &\geq 0_{n_1}. \end{aligned} \tag{II'}$$

We denote the solution sets of (I), (I'), (II), and (II') by X, Z, U, and W, respectively. Unlike (I) and (II), systems (I') and (II') always have solutions, because $0_m \in Z$ and $0_{n+1} \in W$.

Lemma. *Systems* (I) *and* (II) *are not solvable simultaneously.*

Theorems 2 and 4 below imply that there always is a solution to precisely one system, (I) or (II). Therefore, these systems are alternative. The system alternative to (II) reduces to original system (I).

Let us denote the penalty at a point $x \in \Pi_x$ for violating the condition $x \in X$ by pen(x, X). As the penalty, we use the Euclidean norm of the residual vector:

$$\text{pen}(x, X) = [\|(b_1 - A_{11}x_1 - A_{12}x_2)_+\|^2 + \|b_2 - A_{21}x_1 - A_{22}x_2\|^2]^{\frac{1}{2}}.$$

Similarly, we define

$$\text{pen}(u, U) = [\|(A_{11}^{\text{T}}u_1 + A_{21}^{\text{T}}u_2)_+\|^2 + \|A_{12}^{\text{T}}u_1 + A_{22}^{\text{T}}u_2\|^2 + (\rho - b_1^{\text{T}}u_1 - b_2^{\text{T}}u_2)^2]^{\frac{1}{2}}.$$

Here and in what follows, a_+ is the nonnegative part of the vector a; i.e., the ith component of the vector a_+ coincides with the ith component of the vector a if it is nonnegative; otherwise, this component is zero.

Consider the following four problems:

$$I_1 = \min_{x \in \Pi_x} \frac{[\text{pen}(x, X)]^2}{2}, \tag{1}$$

$$I_2 = \min_{u \in \Pi_u} \frac{[\text{pen}(u, U)]^2}{2}, \tag{2}$$

$$I_1^d = \max_{z \in Z} \left\{ \frac{b^{\text{T}}z - \|z\|^2}{2} \right\}, \tag{3}$$

$$I_2^d = \max_{w \in W} \left\{ \frac{\rho w_3 - \|w\|^2}{2} \right\}. \tag{4}$$

The sets Z and W are always nonempty, because they contain the zero vectors. Unlike systems (I) and (II), which may be solvable or not, problems (1)–(4) always have solutions. Moreover, problems (3) and (4) always have unique solutions, because their admissible sets Z and W are nonempty and the strictly concave quadratic objective functions are bounded from above. Problems (1) and (2) are dual to problems (3) and (4), respectively.

We say that a projection of a point $\bar{x}$ on a nonempty closed set X is the point $x^* \in X$ nearest to $\bar{x}$, i.e., such that x^* is a solution to the problem

$$\min_{x \in X} \frac{\|\bar{x} - x\|^2}{2} = \frac{\|\bar{x} - x^*\|^2}{2}.$$

We then write $x^* = \text{pr}(\bar{x}, X)$; the distance from $\bar{x}$ to X is denoted by $\text{dist}(\bar{x}, X) = \|x^* - \bar{x}\|$.

Theorem 1. *Each solution x^* to problem* (1) *determines a unique solution $z^{*\text{T}} = [z_1^{*\text{T}}, z_2^{*\text{T}}]$ to problem* (3) *by the formulas*

$$\begin{aligned} z_1^* &= (b_1 - A_{11}x_1^* - A_{12}x_2^*)_+, \\ z_2^* &= b_2 - A_{21}x_1^* - A_{22}x_2^*, \end{aligned} \tag{5}$$

and the following assertions are valid:

$$\begin{gathered} \|z^*\|^2 = b^{\text{T}}z^*, \\ z^* \perp Ax^*, \quad z^* \perp (b - z^*), \\ z^* = \text{pr}(b, Z), \quad \|z^*\| = \text{pen}(x^*, X), \\ \|b - z^*\| = \text{dist}(b, Z), \\ [\text{pen}(x^*, X)]^2 + [\text{dist}(b, Z)]^2 = \|b\|^2. \end{gathered} \tag{6}$$

Relation (6) follows from the equality of the optimal values of the objective functions for the primal (3) and dual (1) problems. By virtue of (5), this equality is expressed in terms of only z^*, a solution to problem (3).

Theorem 2. *Suppose that $x^{*\text{T}} = [x_1^{*\text{T}}, x_2^{*\text{T}}]$ is an arbitrary solution to problem* (1) *and $z^{*\text{T}} = [z_1^{*\text{T}}, z_2^{*\text{T}}]$ is the unique solution to problem* (3). *Then the following assertions are valid.*

(i) *If* $\|z^*\| = 0$, *then* $I_1 = I_1^d = 0$, $\text{dist}(b, Z) = \|b\|$, *system* (I) *is solvable, and one of its solutions is the vector x^*; system* (II) *is then unsolvable*;

(ii) *If* $\|z^*\| \neq 0$, *then* $I_1 = I_1^d > 0$, $\text{dist}(b, Z) < \|b\|$, *and system* (I) *is unsolvable; system* (II) *is then solvable, and the vector $u^* = \rho z^*/\|z^*\|^2$ is its normal solution.*

Define the $m \times (n+1)$ matrix $\bar{A} = -[A, b]$ and consider the vector $r \in R^{n+1}$ such that $r^{\text{T}} = [0_n^{\text{T}}, \rho]$.

Theorem 3. *Suppose that $u^{*\text{T}} = [u_1^{*\text{T}}, u_2^{*\text{T}}]$ is an arbitrary solution to problem* (2). *Then the solution $w^{*\text{T}} = [w_1^{*\text{T}}, w_2^{*\text{T}}, w_3^*]$ to problem* (4) *is expressed through u^* to problem* (2) *by the formulas*

$$w_1^* = (A_{11}^{\mathrm{T}}u_1^* + A_{21}^{\mathrm{T}}u_2^*)_+,$$
$$w_2^* = A_{12}^{\mathrm{T}}u_1^* + A_{22}^{\mathrm{T}}u_2^*, \quad w_3^* = \rho - b_1^{\mathrm{T}}u_1^* - b_2^{\mathrm{T}}u_2^* \tag{7}$$

and the following assertions are valid:

$$\|w^*\|^2 = \rho w_3^*; \tag{8}$$

$$w^* \perp \bar{A}^{\mathrm{T}}u^*, \quad w^* \perp (r - w^*),$$
$$w^* = \mathrm{pr}(r, W), \quad \|w^*\| = \mathrm{pen}(u^*, U),$$
$$\|r - w^*\| = \mathrm{dist}(r, W),$$
$$[\mathrm{pen}(u^*, U)]^2 + [\mathrm{dist}(r, W)]^2 = \|r\|^2,$$
$$\|w^*\| \le \rho, \quad 0 \le w_3^* \le \rho, \quad \|w_1^*\|^2 + \|w_2^*\|^2 \le \frac{\rho^2}{4}.$$

Relation (8) follows from the equality of the optimal values of the objective functions for the primal (4) and dual (2) problems. By virtue of (7), this equality is expressed in terms of only w^*, a solution to problem (4).

Theorem 4. *Suppose that* $u^{*\mathrm{T}} = [u_1^{*\mathrm{T}}, u_2^{*\mathrm{T}}]$ *is an arbitrary solution to problem* (2) *and* $w^{*\mathrm{T}} = [w_1^{*\mathrm{T}}, w_2^{*\mathrm{T}}, w_3^*]$ *is a solution to problem* (4) *given by formulas* (7). *Then the following assertions are valid.*

(i) *If* $\|w^*\| = 0$, *then* $I_2 = I_2^d = 0$, *system* (II) *is solvable, and one of its solutions is* u^*; *system* (I) *is then unsolvable*;

(ii) *If* $\|w^*\| \ne 0$, *then* $w_3^* > 0$, $I_1 = I_1^d > 0$, *and system* (II) *is unsolvable; system* (I) *is then solvable, and the vector* x^* with the components $x_1^* = \frac{w_1^*}{w_3^*}$, $x_2^* = \frac{w_2^*}{w_3^*}$ *is its solution with minimum Euclidean norm.*

Alternative system (II) admit various representations. As follows from the theorems stated above, the system alternative to (I) is obtained from conjugate system (I') by adding a condition excluding the trivial solution to system (I'). For example, we can require that the solutions to conjugate system (I') satisfy the condition $b^{\mathrm{T}}u > 0$ (as in the Farkas alternative) or the condition $b^{\mathrm{T}}u = 1$ (as in the Gale alternative), or we can impose the nonlinear condition $\|u\|^2 = \rho$, where $\rho > 0$ is an arbitrary fixed value, and so on.

Let us apply the results stated above to linear programming problems. Consider the primal linear programming problem in the form

$$\min_{x \in X} c^{\mathrm{T}}x, \quad X = \{x \in R^n \colon Ax = b, x \ge 0_n\}. \tag{P}$$

Here, A is an $m \times n$ matrix of rank m; $m < n$; $\nu = n - m$ is the defect of the matrix A; and $c, x \in R^n$ and $b \in R^m$ are vectors.

Instead of the traditional necessary and sufficient optimality conditions (1) for linear programming problems, we apply the conditions given in [7]. For this purpose, we introduce a $\nu \times n$ matrix K of rank ν such that $\mathrm{im}K^{\mathrm{T}} = \ker A$, $AK^{\mathrm{T}} = 0$, and $R^n = \mathrm{im}A^{\mathrm{T}} \oplus \mathrm{im}K^{\mathrm{T}}$. Consider $d = Kc \in R^\nu$, the residual vector $v = c - A^{\mathrm{T}}u$, and the affine sets

$$\bar{X} = \{x \in R^n \colon Ax = b\}, \quad \bar{V} = \{v \in R^n \colon Kv = d\}.$$

Let $\bar{x}$ and $\bar{v}$ be arbitrary fixed n-vectors satisfying the conditions $\bar{x} \in \bar{X}$, and $\bar{v} \in \bar{V}$, respectively.

According to [7], the necessary and sufficient minimum conditions for problem (P) have the form

$$\begin{bmatrix} A & 0_{mn} \\ 0_{\nu n} & K \\ \bar{v}^{\mathrm{T}} & \bar{x}^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} = \begin{bmatrix} b \\ d \\ \bar{x}^{\mathrm{T}}\bar{v} \end{bmatrix}, \quad x \ge 0_n, \quad v \ge 0_n. \tag{9}$$

If problem (P) has a solution, then system (9) is consistent, and solving it gives solutions of problem (P) and to the conjugate problem

$$\min_{v \in V} \bar{x}^{\mathrm{T}}v, \quad V = \{v \in R^n \colon Kv = d, v \ge 0_n\}. \tag{C}$$

System (9) comprises $n + 1$ equalities and $2n$ inequalities in $2n$ unknowns. The alternative system has only $n + 1$ unknowns and comprises $2n$ linear inequalities and one equality, namely,

$$\begin{bmatrix} A^{\mathrm{T}} & 0_{n\nu} & \bar{v} \\ 0_{nm} & K^{\mathrm{T}} & \bar{x} \end{bmatrix} \begin{bmatrix} p \\ q \\ \alpha \end{bmatrix} \le 0_{2n}, \quad b^{\mathrm{T}}p + d^{\mathrm{T}}q + \bar{x}^{\mathrm{T}}\bar{v}\alpha = \rho, \tag{10}$$

where $\rho > 0$ is an arbitrary positive constant.

Since system (9) is consistent, alternative system (10) is inconsistent. Problem (2) in this case is written as

$$\min_{p \in R^m} \min_{q \in R^\nu} \min_{\alpha \in R^1} \frac{1}{2}[\|(A^{\mathrm{T}}p + \bar{v}\alpha)_+\|^2 + \|(K^{\mathrm{T}}q + \bar{x}\alpha)_+\|^2$$
$$+ (\rho - b^{\mathrm{T}}p - d^{\mathrm{T}}q - \bar{x}^{\mathrm{T}}\bar{v}\alpha)^2].$$

Solving this problem yields optimal vectors p^*, q^*, and α^*, which determine the discrepancies of inconsistent system (10) according to

$$w_x^* = (A^{\mathrm{T}}p^* + \bar{v}\alpha^*)_+, \; w_v^* = (K^{\mathrm{T}}q^* + \bar{x}\alpha^*)_+,$$
$$w_3^* = \rho - b^{\mathrm{T}}p^* - d^{\mathrm{T}}q^* - \bar{x}^{\mathrm{T}}\bar{v}\alpha^*.$$

By Theorem 4, normal solutions to system (9) are given by the formulas

$$\tilde{x}^* = \frac{w_x^*}{w_3^*}, \quad \tilde{v}^* = \frac{w_v^*}{w_3^*},$$

and they are simultaneously normal solutions to problems (P) and (C).

Thus, we have reduced solving a linear programming problem to once unconstrained minimization of a convex differentiable piecewise quadratic function of $n + 1$ variables.

ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic Research (project no. 01-01-00804) and the state program of support of leading scientific schools (project no. 00-15-96080).

REFERENCES

1. Chernikov, S.N., *Lineinye neravenstva* (Linear Inequalities), Moscow: Nauka, 1968.
2. Eremin, I.I., Mazurov, V.D., and Astaf'ev, N.N., *Nesobstvennye zadachi lineinogo i vypuklogo programmirovaniya* (Improper Problems of Linear and Convex Programming), Moscow: Nauka, 1983.
3. Razumikhin, B.S., *Fizicheskie modeli i metody teorii ravnovesiya v programmirovanii i ekonomike* (Physical Models and the Methods of Equilibrium Theory in Programming and Economics), Moscow: Nauka, 1975.
4. Gale, D., *The Theory of Linear Economic Models*, New York, 1960. Translated under the title *Teoriya lineinykh ekonomicheskikh modelei,* Moscow: Inostrannaya Literatura, 1963.
5. Mangasarian, O.L., *Nonlinear Programming*, Philadelphia: SIAM, 1994.
6. Giannessi, F., *Encyclopedia of Optimization,* Dordrecht: Klüwer, vol. 5, pp. 437–444.
7. Golikov, A.I. and Evtushenko, Yu.G., *Zh. Vych. Mat. Mat. Fiz.*, vol. 40, no. 12, pp. 1766–1786.