SIAM REV.
Vol. 39, No. 3, pp. 503–507, September 1997
© 1997 Society for Industrial and Applied Mathematics
008

AN ELEMENTARY PROOF OF FARKAS' LEMMA*

ACHIYA DAX†

Abstract. Farkas' lemma is one of the key results in optimization. Yet, it is not a trivial conclusion, and its proof contains certain difficulties. In this note we propose a new proof which is based on elementary arguments.

Key words. Farkas' lemma, elementary proof, active set method

AMS subject classifications. 65K05, 65K10

PII. S0036144594295502

Farkas' lemma is a classical result, first published in 1902. It belongs to a class of statements called "theorems of the alternative," which characterizes the optimality conditions of several problems. A proof of Farkas' lemma can be found in almost any optimization textbook. See, for example, [1–11]. Early proofs of this observation are rather formal and don't make clear why the theorem works; see, e.g., [5, p. 44]. Recent proofs are usually based on projection (separation) theorems. This approach has a simple geometrical interpretation and is far more intuitive. Nevertheless, Farkas' lemma is still considered a "pedagogical annoyance" because some parts of it are easy to verify while the main result cannot be proved in an elementary way. For example, see [6, p. 295]. In this note we attempt to overcome this difficulty by presenting a simple self-contained proof which is based on elementary arguments.

Let A be a real $m \times n$ matrix and let $\mathbf{c}$ be a real nonzero n-vector. Then Farkas' lemma says that either the primal system

$$A\mathbf{x} \geq \mathbf{0} \quad \text{and} \quad \mathbf{c}^T\mathbf{x} < 0 \tag{1}$$

has a solution $\mathbf{x} \in \mathbb{R}^n$ or the dual system

$$A^T\mathbf{y} = \mathbf{c} \quad \text{and} \quad \mathbf{y} \geq \mathbf{0} \tag{2}$$

has a solution $\mathbf{y} \in \mathbb{R}^m$, but never both. Note that $\mathbf{c}^T\mathbf{x} < 0$ implies $\mathbf{x} \neq \mathbf{0}$, while $A^T\mathbf{y} = \mathbf{c}$ means $\mathbf{y} \neq \mathbf{0}$. (The dimension of the null vector, $\mathbf{0}$, depends on the context.) If both (1) and (2) hold then the inequality

$$\mathbf{c}^T\mathbf{x} = (A^T\mathbf{y})^T\mathbf{x} = \mathbf{y}^T A\mathbf{x} \geq 0$$

contradicts $\mathbf{c}^T\mathbf{x} < 0$. Hence, it is not possible that both systems are solvable. The question of which of the two systems is solvable is answered by considering the bounded least squares problem

$$\text{minimize } ||A^T\mathbf{y} - \mathbf{c}||^2 \tag{3a}$$

$$\text{subject to } \mathbf{y} \geq \mathbf{0}, \tag{3b}$$

where $||\ ||$ denotes the Euclidean norm. Let $\mathbf{y}^*$ be a given point in $\mathbb{R}^m$, and let

$$\mathbf{r}^* = A^T\mathbf{y}^* - \mathbf{c} \tag{4}$$

*Received by the editors July 5, 1994; accepted for publication November 22, 1995.
http://www.siam.org/journals/sirev/39-3/29550.html
†Hydrological Service, P.O.B. 6381, Jerusalem 91063, Israel (dax@vms.huji.ac.il).

denote the corresponding residual vector. Then, by Lemma 2 below, $\mathbf{y}^*$ solves (3) if and only if $\mathbf{y}^*$ and $\mathbf{r}^*$ satisfy the conditions

$$\mathbf{y}^* \geq \mathbf{0}, \quad A\mathbf{r}^* \geq \mathbf{0}, \quad \text{and} \quad (\mathbf{y}^*)^T A\mathbf{r}^* = 0. \tag{5}$$

These conditions will be used later to establish the existence of a point $\mathbf{y}^* \in \mathbb{R}^m$ that solves (3). Combining (4) and (5) gives

$$\mathbf{c}^T\mathbf{r}^* = (A^T\mathbf{y}^* - \mathbf{r}^*)^T\mathbf{r}^* = (\mathbf{y}^*)^T A\mathbf{r}^* - (\mathbf{r}^*)^T\mathbf{r}^* = -||\mathbf{r}^*||^2,$$

which leads to the following conclusion.

THEOREM 1. *Let $\mathbf{y}^*$ solve* (3) *and let $\mathbf{r}^* = A^T\mathbf{y}^* - \mathbf{c}$ denote the corresponding residual vector. If $\mathbf{r}^* = \mathbf{0}$ then $\mathbf{y}^*$ solves* (2). *Otherwise, $\mathbf{r}^*$ solves* (1) *and $\mathbf{c}^T\mathbf{r}^* = -||\mathbf{r}^*||^2$.*

LEMMA 2. *Let $\mathbf{y}^* = (y_1^*, \ldots, y_m^*)^T$ be a given point in $\mathbb{R}^m$ and let $\mathbf{r}^*$ be defined by* (4). *Then $\mathbf{y}^*$ solves* (3) *if and only if $\mathbf{y}^*$ and $\mathbf{r}^*$ satisfy* (5).

Proof. Assume that $\mathbf{y}^*$ solves (3) and consider the one-parameter quadratic functions

$$f_i(\theta) = ||A^T(\mathbf{y}^* + \theta\mathbf{e}_i) - \mathbf{c}||^2 = ||\theta\mathbf{a}_i + \mathbf{r}^*||^2, \quad i = 1, \ldots, m,$$

where $\mathbf{a}_i^T$ denotes the ith row of A, θ is a real variable, and $\mathbf{e}_i$ denotes the ith column of the $m \times m$ unit matrix. Then, clearly, $\theta = 0$ solves the problem

$$\begin{aligned} &\text{minimize } f_i(\theta) \\ &\text{subject to } \ y_i^* + \theta \geq 0. \end{aligned}$$

Therefore, since $f_i'(0) = 2\mathbf{a}_i^T\mathbf{r}^*$,

$$y_i^* > 0 \text{ implies } \ \mathbf{a}_i^T\mathbf{r}^* = 0, \text{ while } \ y_i^* = 0 \text{ implies } \ \mathbf{a}_i^T\mathbf{r}^* \geq 0,$$

which constitutes (5).

Conversely, assume that (5) holds and let $\mathbf{z}$ be an arbitrary point in $\mathbb{R}^m$ such that $\mathbf{z} \geq \mathbf{0}$. Let the m-vector $\mathbf{u} = (u_1, \ldots, u_m)^T$ be obtained from $\mathbf{z}$ by the rule $\mathbf{u} = \mathbf{z} - \mathbf{y}^*$. Then $y_i^* = 0$ implies $u_i \geq 0$, while (5) leads to

$$\mathbf{u}^T A\mathbf{r}^* \geq 0.$$

Hence, the identity

$$||A^T\mathbf{z} - \mathbf{c}||^2 = ||A^T\mathbf{y}^* - \mathbf{c}||^2 + 2\mathbf{u}^T A\mathbf{r}^* + ||A^T\mathbf{u}||^2$$

shows that

$$||A^T\mathbf{z} - \mathbf{c}||^2 \geq ||A^T\mathbf{y}^* - \mathbf{c}||^2. \qquad \square$$

It is left to establish the existence of a point $\mathbf{y}^*$ that solves (3). This aim is achieved by introducing a simple iterative algorithm whose kth iteration, $k = 1, 2, \ldots$, consists of the following two steps.

Step 1: *Solving an unconstrained least squares problem.* Let $\mathbf{y}_k = (y_1, \ldots, y_m)^T \geq \mathbf{0}$ denote the current estimate of the solution at the beginning of the kth iteration. Define

$$\mathbf{r}_k = A^T\mathbf{y}_k - \mathbf{c}, \quad V_k = \{i \mid \ y_i = 0\}, \quad \text{and} \quad W_k = \{i \mid \ y_i > 0\}.$$

The number of indices in W_k is denoted by t. If $t = 0$ or $\mathbf{r}_k = \mathbf{0}$ then skip to Step 2. Let A_k be the $t \times n$ matrix whose rows are $\mathbf{a}_i^T$, $i \in W_k$. The order of the rows does not matter. Hence, it is assumed for simplicity that

$$A_k^T = [\mathbf{a}_1, \dots, \mathbf{a}_t], \quad W_k = \{1, \dots, t\}, \quad \text{and} \quad V_k = \{t+1, \dots, m\}.$$

Let the t-vector $\mathbf{w}_k = (w_1, \dots, w_t)^T$ solve the unconstrained least squares problem

$$\text{minimize} \ \ ||A_k^T \mathbf{w} - \mathbf{r}_k||^2.$$

Observe that $\mathbf{0}$ solves this problem if and only if $A_k \mathbf{r}_k = \mathbf{0}$. In this case, skip to Step 2. Otherwise, define a nonzero search direction $\mathbf{u}_k = (u_1, \dots, u_m)^T \in \mathbb{R}^m$ by the following rule:

$$u_i = w_i \ \text{ for } \ i = 1, \dots, t \ \text{ and } \ u_i = 0 \ \text{ for } \ i = t+1, \dots, m.$$

Then the next point is defined as

$$\mathbf{y}_{k+1} = \mathbf{y}_k + \theta_k \mathbf{u}_k,$$

where $\theta_k > 0$ is the largest number in the interval $[0, 1]$ that keeps the point $\mathbf{y}_k + \theta \mathbf{u}_k$ feasible. In other words, θ_k is the smallest number in the set $\{1\} \cup \{-y_i/u_i \mid \ u_i < 0\}$.

Step 2: *Testing optimality and moving away from a dead point.* Here $A_k \mathbf{r}_k = \mathbf{0}$, which means that $\mathbf{y}_k$ solves the problem

$$\text{minimize} \ \ ||A^T \mathbf{y} - \mathbf{c}||^2 \tag{6a}$$

$$\text{subject to} \ \ y_i = 0 \ \text{ for } \ i \in V_k \tag{6b}$$

$$\text{and} \ \ y_i \geq 0 \ \text{ for } \ i \in W_k. \tag{6c}$$

In this case, $\mathbf{y}_k$ is called a "dead point." To test whether or not $\mathbf{y}_k$ is optimal, we compute an index j such that

$$\mathbf{a}_j^T \mathbf{r}_k = \min\{\mathbf{a}_i^T \mathbf{r}_k \mid \ i \in V_k\}.$$

If $\mathbf{a}_j^T \mathbf{r}_k \geq 0$ then $\mathbf{y}_k$ and $\mathbf{r}_k$ satisfy (5). Therefore, by Lemma 2, $\mathbf{y}_k$ solves (3) and the algorithm terminates. Otherwise, the next point is defined as

$$\mathbf{y}_{k+1} = \mathbf{y}_k - (\mathbf{a}_j^T \mathbf{r}_k / \mathbf{a}_j^T \mathbf{a}_j) \mathbf{e}_j,$$

where $\mathbf{e}_j$ denotes the jth column of the $m \times m$ unit matrix. Note that $-\mathbf{a}_j^T \mathbf{r}_k / \mathbf{a}_j^T \mathbf{a}_j > 0$, and this point solves the problem

$$\text{minimize} \ \ f(\theta) = ||A^T(\mathbf{y}_k + \theta \mathbf{e}_j) - \mathbf{c}||^2.$$

The finite termination of the above algorithm is a consequence of the following properties:

(a) The objective function is strictly decreasing at each iteration, that is,

$$||A^T \mathbf{y}_k - \mathbf{c}||^2 > ||A^T \mathbf{y}_{k+1} - \mathbf{c}||^2.$$

(b) If $\theta_k = 1$ then $\mathbf{y}_{k+1}$ is a dead point. Otherwise, when $0 < \theta_k < 1$, t decreases. Hence, it is not possible to perform more than m iterations without reaching a dead point.

(c) Each time we reach a dead point, the current point solves (6).

(d) There is a finite number of such problems. Yet, because of (a), it is not possible to "visit" the same problem twice.

It is interesting to compare the new proof with other approaches. A traditional way for proving the existence of a point $\mathbf{y}^*$ which solves (3) is based on the observation that

$$Z = \{A^T\mathbf{y} \mid \ \mathbf{y} \geq \mathbf{0}\}$$

is a closed set of $\mathbb{R}^n$. However, contrary to the impression that a hasty analysis might leave, proving this assertion indeed requires elaboration. See, for example, [1, pp. 332–334] or [9, p. 10]. Using the closure of Z, we obtain that

$$\mathbb{B} = \{\mathbf{z} \mid \ \mathbf{z} \in Z \ \text{ and } \ ||\mathbf{z} - \mathbf{c}|| \leq ||\mathbf{c}||\}$$

is a nonempty closed bounded set of $\mathbb{R}^n$. Note also that $\varphi(\mathbf{x}) = ||\mathbf{x} - \mathbf{c}||^2$ is a continuous function of $\mathbf{x}$. Therefore, by the well-known Weirestrass' theorem, $\varphi(\mathbf{x})$ achieves its minimum over $\mathbb{B}$. Let $\mathbf{z}^* \in \mathbb{B}$ denote the minimizing point. Since $\mathbb{B} \subseteq Z$, there exists a point $\mathbf{y}^* \in \mathbb{R}^m$ such that $\mathbf{y}^* \geq \mathbf{0}$ and $\mathbf{z}^* = A^T\mathbf{y}^*$; therefore, $\mathbf{y}^*$ solves (3).

The point $\mathbf{z}^* = A^T\mathbf{y}^*$ is the projection of $\mathbf{c}$ on Z. Since Z is a convex set and $\varphi(\mathbf{x})$ is a strictly convex function, $\mathbf{z}^*$ is unique. However, $\mathbf{y}^*$ is not necessarily unique. It is also worthwhile mentioning a relatively less well known feature of $\mathbf{r}^*$, due to Powell [10].

COROLLARY 3. *Let $\mathbf{y}^*$ and $\mathbf{r}^*$ be as in Theorem* 1 *and assume that $\mathbf{r}^* \neq \mathbf{0}$. Then the vector $\mathbf{r}^*/||\mathbf{r}^*||$ solves the steepest descent problem*

$$\text{minimize} \ \ \mathbf{c}^T\mathbf{x} \tag{7a}$$

$$\text{subject to} \ \ A\mathbf{x} \geq \mathbf{0} \text{ and } ||\mathbf{x}|| = 1. \tag{7b}$$

Proof. Let $\mathbf{x}$ satisfy the constraints (7b). Then

$$(\mathbf{y}^*)^T A\mathbf{x} \geq 0,$$

while the Cauchy–Schwartz inequality gives

$$|(\mathbf{r}^*)^T\mathbf{x}| \leq ||\mathbf{r}^*|| \cdot ||\mathbf{x}|| = ||\mathbf{r}^*||.$$

Combining these relations shows that

$$\mathbf{c}^T\mathbf{x} = (A^T\mathbf{y}^* - \mathbf{r}^*)^T\mathbf{x} = (\mathbf{y}^*)^T A\mathbf{x} - (\mathbf{r}^*)^T\mathbf{x} \geq -(\mathbf{r}^*)^T\mathbf{x} \geq -|(\mathbf{r}^*)^T\mathbf{x}| \geq -||\mathbf{r}^*||.$$

Therefore, since $\mathbf{c}^T\mathbf{r}^*/||\mathbf{r}^*|| = -||\mathbf{r}^*||$, the claim is proved. $\square$

Concluding remarks. A common way of proving Farkas' lemma is to apply the separating hyperplane theorem with projection of $\mathbf{c}$ on Z. See, for example, [1, 4, 6, 8, 9, 10, 11]. This approach requires establishing both the separating hyperplane theorem and the closure of Z. In our proof, Theorem 1 replaces the separating hyperplane theorem, while $\mathbf{y}^*$ is obtained via a typical "active set" method.

Other theorems of the alternative are also related to bounded least squares problems and steepest descent directions. Hence, a similar method of proof can be used to establish the other theorems. The ability to compute the steepest descent direction provides an effective way for resolving degeneracy in certain active set methods. The reader is referred to [2] for a detailed discussion of these issues.

REFERENCES

[1] P.G. Ciarlet, *Introduction to Numerical Linear Algebra and Optimization*, Cambridge University Press, Cambridge, 1989.

[2] A. Dax, *The relationship between theorems of the alternative, least norm problems, steepest descent directions, and degeneracy: A review*, Ann. Oper. Res., 46 (1993), pp. 11–60.

[3] J. Farkas, *Über die Theorie der einfachen Ungleichungen*, J. Reine Angew. Math., 124 (1902), pp. 1–24.

[4] R. Fletcher, *Practical Methods of Optimization, Vol. 2: Constrained Optimization*, John Wiley, New York, 1981.

[5] D. Gale, *The Theory of Linear Economic Models*, McGraw–Hill, New York, 1960.

[6] P.E. Gill, W. Murray, and M.H. Wright, *Numerical Linear Algebra and Optimization*, Vol. 1, Addison–Wesley, Reading, MA, 1991.

[7] O.L. Mangasarian, *Nonlinear Programming*, McGraw–Hill, New York, 1969.

[8] G.P. McCormick, *Nonlinear Programming*, John Wiley, New York, 1983.

[9] M.R. Osborne, *Finite Algorithms in Optimization and Data Analysis*, John Wiley & Sons, Chichester, 1985.

[10] M.J.D. Powell, *Introduction to constrained optimization*, in Numerical Methods for Constrained Optimization, P.E. Gill and W. Murray, eds., Academic Press, New York, 1974, pp. 1–28.

[11] G. Zoutendijk, *Methods of Feasible Directions*, Elsevier–North Holland, Amsterdam, 1960.