Mathematica Pannonica
19/1 (2008), 117–124

THE SECTOROID VERSION OF THE FARKAS LEMMA

Zoltán **Kánnai**

Department of Mathematics, Corvinus University Budapest, H-1828 Budapest, Hungary

Received: April 2008

MSC 2000: 90 C 30

Keywords: Farkas' lemma, elementary proof, nonlinear programming.

Abstract: The following statement for linear inequalities has fundamental importance in mathematical programming:
Farkas' Lemma: *For given vectors* $a, a_1, \dots, a_m \in \mathbb{R}^n$*, if the inequality* $\langle a, x\rangle \leq 0$ *is a consequence of the system of inequalities* $\langle a_1, x\rangle \leq \leq 0, \dots, \langle a_m, x\rangle \leq 0$*, then* a *is a nonnegative combination of* $a_1, \dots, a_m$.
This theorem is generally known as a relatively hard statement. The simplification of its proof has mainly educational importance. The original proof of Farkas [2], and also further works (e.g., [3]) are quite elementary, but not suggestive at all; other papers (e.g. [5]) are more elegant but use hard techniques of Euclidean or even Hilbert spaces. Recently, Broyden [1], and Komornik [4] gave simpler proofs, but both of them still capitalize the Euclidean metric property. This paper presents Farkas' Lemma as one of the introductory statements of linear algebra, suitable to discuss even at the beginning of the linear algebra course. Moreover a sectoroid version of Farkas' Lemma is also proved. The sectoroid extension of the Farkas Lemma makes possible to discuss smooth programming problems satisfying a sectoroid type regularity condition, without hard devices of functional analysis or implicit function theorem. That problem in itself is important in several applications, especially in optimization problems with budget type constraints.

1. Introduction

The fundamental idea is the following: take a homogenous linear inequality system

$$(*)\qquad \begin{cases} a_{1,1}\cdot x_1 + a_{1,2}\cdot x_2 + \dots + a_{1,n}\cdot x_n & \le & 0 \\ a_{2,1}\cdot x_1 + a_{2,2}\cdot x_2 + \dots + a_{2,n}\cdot x_n & \le & 0 \\ \vdots \qquad \ddots & & \vdots \\ \vdots \qquad\qquad \dots \qquad \dots & & \vdots \\ a_{m,1}\cdot x_1 + a_{m,2}\cdot x_2 + \dots + a_{m,n}\cdot x_n & \le & 0 \end{cases}$$

of real coefficients, and the linear inequality

$$(**)\qquad b_1\cdot x_1 + b_2\cdot x_2 + \dots + b_n\cdot x_n \quad \le \quad 0.$$

It is well-known that a nonnegative combination of the inequalities from $(*)$ holds true for all arguments satisfying $(*)$. That is, if $(**)$ comes from $(*)$ in this way, then the solution set of $(*)$ is contained in the solution set of $(**)$.

Much less obvious is the inverse statement, i.e. if the solution set of $(*)$ is contained in the solution set of $(**)$, then $(**)$ is a nonnegative combination of $(*)$ indeed. This is the famous lemma proven by Gyula Farkas [2] used at first in mechanics, but nowadays it has fundamental importance in the game theory. This famous result became a principal reference in Optimization Theory after the publication of the paper of Kuhn and Tucker [5]. In that paper Farkas' fundamental theorem on linear inequalities was used to derive necessary condition for optimality for the nonlinear programming problem. The results obtained led to a rapid development of nonlinear optimization theory. Because of its historical importance and its many applications in convex analysis, duality theory, optimality conditions, etc., the Farkas Lemma can be considered a cornerstone in optimization.

Farkas' merit was not only to have provided a rigorous proof of a result on linear inequality system, but the invention of the Karush–Kuhn–Tucker necessary optimality condition itself (in case of regularity had been taken granted).

A true and detailed evaluation of Gyula Farkas' work is due to András Prékopa, who made significant attempts in order to make known within the wide international scientific community that Gyula Farkas might be considered not merely as the inventor of the Farkas Lemma, but as one of the forerunner of the modern optimization theory (including equilibrium problems and variational inequalities) (cf. [6] and [7]).

Since there are many proofs for this lemma. These are mostly either too complicated, or unnatural, or, they use relatively hard devices. In our opinion the key for an elementary, short and natural proof is the specific language of the abstract linear algebra.

Namely, at the beginning of this topic, one meets the following statement: if X is a vector space over the field $\mathbb{F}$, and the linear functionals $\varphi, \varphi_1, \varphi_2, \ldots, \varphi_m$ of the dual space X' satisfy that

$$\bigcap_{k=1}^{m} \varphi_k^{-1}(0) \subseteq \varphi^{-1}(0),$$

then φ is a linear combination of $\varphi_1, \varphi_2, \ldots, \varphi_m$. We will refer only to the case $m = 1$:

$$\varphi_1^{-1}(0) \subseteq \varphi^{-1}(0) \Gamma\!\Rightarrow \exists c : \varphi = c \cdot \varphi_1 .$$

Farkas' Lemma is obviously the "nonnegative version" of this statement (with $\varphi = \langle a, \cdot \rangle$, $\varphi_k = \langle a_k, \cdot \rangle$; $k = 1, \ldots, m$), and has a similarly simple proof, with no Euclidean devices.

2. Farkas' Lemma and its sectoroid version

Denote $\mathbb{R}_- := (-\infty, 0]$ and $\mathbb{R}_-^{o} := (-\infty, 0)$.

Theorem 1 (Farkas)**.** *Let X be a real vector space and $\varphi, \varphi_1, \varphi_2, \ldots, \varphi_m \in$ $\in X'$ linear functionals. If*

$$\bigcap_{k=1}^{m} \varphi_k^{-1}(\mathbb{R}_-) \subseteq \varphi^{-1}(\mathbb{R}_-),$$

then φ is a nonnegative combination of $\varphi_1, \varphi_2, \ldots, \varphi_m$.

Proof. We prove by induction with respect to m.

$m = 1$: Let X be a real vector space and $\varphi, \varphi_1 \in X'$ for which $\varphi_1^{-1}(\mathbb{R}_-) \subseteq \varphi^{-1}(\mathbb{R}_-)$. Hence for every $x \in \ker \varphi_1$ we get $\varphi x \leq 0$, and also $\varphi(-x) \leq 0$, thus, $\varphi x = 0$, i.e., $\varphi_1^{-1}(0) \subseteq \varphi^{-1}(0)$. Then $\varphi = c \cdot \varphi_1$ ($c \in \mathbb{R}$). Now if $\varphi_1 = \mathbf{0}$, then change c to 0. Otherwise, there exists a vector $v \in X$ with $\varphi_1 v < 0$. Then by the condition we have $\varphi v \leq 0$, and

$$c = \frac{\varphi v}{\varphi_1 v} \geq 0.$$

Thus, for $m = 1$ the theorem is true.

Now assume that the statement is true for m. Take a real vector space X and linear functionals $\varphi, \varphi_1, \varphi_2, \dots, \varphi_{m+1} \in X'$ such that

$$\bigcap_{k=1}^{m+1} \varphi_k^{-1}(\mathbb{R}_-) \subseteq \varphi^{-1}(\mathbb{R}_-).$$

If $\bigcap\limits_{k=1}^{m} \varphi_k^{-1}(\mathbb{R}_-) \subseteq \varphi^{-1}(\mathbb{R}_-)$ also holds true, then we are ready by the assumption. In the opposite case, there is a vector $v \in X$ such that $\varphi_1 v, \dots, \varphi_m v \leq 0$, but $\varphi v > 0$, and of course, $\varphi_{m+1} v > 0$. Then put $L := \varphi_{m+1}^{-1}(0)$, and apply the induction assumption to the restricted functionals $\varphi \mid_L, \varphi_1 \mid_L, \dots, \varphi_m \mid_L$. Obviously

$$\bigcap_{k=1}^{m} (\varphi_k \mid_L)^{-1}(\mathbb{R}_-) = L \cap \bigcap_{k=1}^{m} \varphi_k^{-1}(\mathbb{R}_-) = L \cap \bigcap_{k=1}^{m+1} \varphi_k^{-1}(\mathbb{R}_-) \subseteq$$
$$\subseteq L \cap \varphi^{-1}(\mathbb{R}_-) = (\varphi \mid_L)^{-1}(\mathbb{R}_-),$$

hence by the induction assumption there are nonnegative constants $\lambda_1, \lambda_2, \dots, \lambda_m$ such that $\varphi \mid_L = \sum\limits_{k=1}^{m} \lambda_k \cdot \varphi_k \mid_L$. Therefore $\left(\varphi - \sum\limits_{k=1}^{m} \lambda_k \varphi_k\right) \mid_L$ is identically zero, implying that

$$\ker \varphi_{m+1} = L \subseteq \ker \left(\varphi - \sum_{k=1}^{m} \lambda_k \varphi_k\right).$$

Again, there is a constant $c \in \mathbb{R}$ such that

$$\varphi - \sum_{k=1}^{m} \lambda_k \varphi_k = c \cdot \varphi_{m+1}.$$

Consequently $\varphi v - \sum\limits_{k=1}^{m} \lambda_k \varphi_k v = c \cdot \varphi_{m+1} v$. Since $\varphi v > 0 \geq \varphi_1 v, \dots, \varphi_m v$, the left-hand side is positive and $\varphi_{m+1} v > 0$, thus we get $c > 0$, and

$$\varphi = \sum_{k=1}^{m} \lambda_k \varphi_k + c\varphi_{m+1}$$

is a combination with nonnegative weights. Hence the statement is true for $m + 1$, by which the proof is completed. ◊

Corollary 2. *Let E be a Euclidean space and $C = \mathrm{cone}(a_1, a_2, \dots, a_m) \subseteq E$ be a finite convex cone. For arbitrary $a \in C^{--} = \{a_1, a_2, \dots, a_m\}^{--}$, applying Farkas' Lemma for the functionals $\langle \cdot \mid a \rangle, \langle \cdot \mid a_1 \rangle, \dots, \langle \cdot \mid a_m \rangle$ we just get that $a \in C$, hence $C^{--} = C$ for every finite convex cone C. This is one of the most popular form of Farkas' Lemma.*

In nonlinear programming there is a very often used regularity condition of geometric form, which we reformulate to an algebraic one, for the sake of fluent discussion.

Definition 1. The system of the vectors $x_1, x_2, \ldots, x_n \in X$ we call **sectoroid** if any nontrivial nonnegative combination of them differs form zero:

$$\alpha_1, \alpha_2, \ldots, \alpha_n \geq 0,\ \sum_{k=1}^{n} \alpha_k x_k = \mathbf{0}_X \ \Longrightarrow\ \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0.$$

Sectoroidity is certainly a much weaker property than linear independence. For example, in the plain arbitrarily many nonzero vectors lying in a fixed acute angular domain are sectoroid. Note that vectors $x_1, x_2, \ldots, x_n \in X$ are sectoroid if and only if their convex hull does not contain the origin. This property for nonzero vectors is also equivalent to the following: their convex cone hull does not contain a complete line. Now we show that sectoroidity is equivalent to the usual geometric regularity condition at issue.

Lemma 1. *The functionals $\varphi_1, \varphi_2, \ldots, \varphi_m \in X'$ are sectoroid if and only if* $\bigcap_{k=1}^{m} \varphi_k^{-1}\left(\mathbb{R}_-^{\circ}\right) \neq \emptyset$.

Proof. Suppose that $\varphi_1, \varphi_2, \ldots, \varphi_m \in X'$ are sectoroid. At first note that for every index $1 \leq p \leq m$,

$$\bigcap_{k=1}^{m} \varphi_k^{-1}\left(\mathbb{R}_-\right) \nsubseteq \varphi_p^{-1}\left(0\right),$$

because in opposite case we have that

$$\bigcap_{k=1}^{m} \varphi_k^{-1}\left(\mathbb{R}_-\right) \subseteq \varphi_p^{-1}\left(0\right) \subseteq \left(-\varphi_p\right)\left(\mathbb{R}_-\right),$$

so by Farkas' Lemma $-\varphi_p$ would be contained in the convex cone hull of $\varphi_1, \varphi_2, \ldots, \varphi_m$, which is contradiction with sectoroidity. Hence for every $1 \leq p \leq m$, there is a vector

$$v_p \in \bigcap_{k=1}^{m} \varphi_k^{-1}\left(\mathbb{R}_-\right) \setminus \varphi_p^{-1}\left(0\right).$$

Since for all $1 \leq k, p \leq m$ we have $\varphi_k v_p \leq 0$ and $\varphi_k v_k < 0$, hence by $w := v_1 + \cdots + v_m$ we get $\varphi_k w < 0$ for every k, i.e.

$$w \in \bigcap_{k=1}^{m} \varphi_k^{-1}\left(\mathbb{R}_-^{\circ}\right).$$

Conversely, suppose that there is a vector $w \in \bigcap_{k=1}^{m} \varphi_k^{-1}\left(\mathbb{R}_-^{\mathrm{o}}\right)$. Now if $\sum_{k=1}^{n} \alpha_k \varphi_k = \mathbf{0}_{X'}$ with nonnegative weights, then by every $\alpha_k \varphi_k w$ being nonpositive, and by $\sum_{k=1}^{n} \alpha_k \varphi_k w = 0$ we get that every $\alpha_k \varphi_k w$ is 0, hence $\alpha_k = 0$. So $\varphi_1, \varphi_2, \ldots, \varphi_m \in X'$ are sectoroid. ◊

Now turn to the sectoroid version of Farkas' Lemma.

Theorem 3. *Let $\varphi_1, \varphi_2, \ldots, \varphi_m \in X'$ be sectoroid functionals. If for an additional functional $\varphi \in X'$,*

$$\bigcap_{k=1}^{m} \varphi_k^{-1}\left(\mathbb{R}_-^{\mathrm{o}}\right) \subseteq \varphi^{-1}\left(\mathbb{R}_-\right),$$

then φ is a nonnegative combination of $\varphi_1, \varphi_2, \ldots, \varphi_m$.

Proof. We need check only the condition of Farkas' Lemma. By the previous lemma we have a vector $w \in \bigcap_{k=1}^{m} \varphi_k^{-1}((-\infty, 0))$. Now for any $x \in \bigcap_{k=1}^{m} \varphi_k^{-1}(\mathbb{R}_-)$ and $t > 0$ we know that

$$\varphi_k(x + tw) = \varphi_k x + t \cdot \varphi_k w < 0,$$

for every k, thus,

$$x + tw \in \bigcap_{k=1}^{m} \varphi_k^{-1}\left(\mathbb{R}_-^{\mathrm{o}}\right) \subseteq \varphi^{-1}((-\infty, 0]),$$

i.e. for every k we have that $\varphi x + t \cdot \varphi w = \varphi(x + tw) \leq 0$, whence by tending $t \to 0$ we get $\varphi x \leq 0$. By this we have shown that

$$\bigcap_{k=1}^{m} \varphi_k^{-1}((-\infty, 0]) \subseteq \varphi^{-1}((-\infty, 0]). \quad \Diamond$$

3. Smooth programming

Remark 1. Suppose that the function $g : \mathbb{R} \to \mathbb{R}$ is differentiable at 0, moreover $g'(0) < 0$. Then in a suitable right neighbourhood $(0, \delta)$, the values taken by g are less than $g(0)$.

Let X be a real normed space, $G \subseteq X$ be an open set, $m_1, m_2, \ldots$ $\ldots, m_n \in \mathbb{R}$, moreover $f, f_1, \ldots, f_n : G \to \mathbb{R}$ be given functions. Consider the constrained optimalization problem

$$(1)\qquad \begin{cases} f(x) \to \max \\ x \in G \\ f_1(x) \le m_1,\ f_2(x) \le m_2, \cdots,\ f_n(x) \le m_n \end{cases}.$$

Theorem 4. *Suppose that $a \in G$ is a solution to problem (1), moreover the functions $f, f_1, \ldots, f_n$ are differentiable at a, and the system of the functionals*

$$\left(f_k'(a) : 1 \le k \le n,\ f_k(a) = m_k\right)$$

is sectoroid. Then there are nonnegative numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that

$$f'(a) = \sum_{k=1}^{n} \lambda_k f_k'(a),$$

furthermore $\lambda_k \cdot (f_k(a) - m_k) = 0$ for every $1 \le k \le m$ (complementary conditions).

Proof. Denote by A the set of indices k fulfilling $f_k(a) = m_k$, and by B the set of indices k fulfilling $f_k(a) < m_k$. Take $v \in \bigcap_{k \in A} [f_k'(a)]^{-1}((-\infty, 0))$. Then for every index $k \in A$ we have $[f_k'(a)]\, v < 0$, i.e., the function $f_{k,a,v} :$ $: t \mapsto f_k(a + tv)$ satisfies $f_{k,a,v}'(0) < 0$, and of course, $f_{k,a,v}(0) = f_k(a)$. So for every $k \in A$ by the above remark there is a right neighborhood U_k of 0 such that for arbitrary $t \in U_k$, the value $f_k(a + tv) = f_{k,a,v}(t)$ is less than $f_k(a) = m_k$. Hence $H_1 = \cap_{k \in A} U_k$ is a right neighborhood of 0 such that for arbitrary $t \in H_1$ and $k \in A$ we have $f_k(a + tv) < m_k$.

On the other hand, for every $k \in B$, by the continuity of f_k at a, and by $f_k(a) < m_k$ we obtain that $f_k(a + tv) < m_k$ in a suitable right neighborhood V_k of 0. Hence $H_2 = \cap_{k \in B} V_k$ is a right neighborhood of 0 such that for arbitrary $t \in H_2$ and $k \in B$ we have $f_k(a + tv) < m_k$.

Now for every $t \in H_1 \cap H_2$ we have $a + tv \in G$, furthermore

$$f_k(a + tv) < m_k$$

for every $1 \le k \le n$, thus, by a being a solution to (1), we obtain $f(a) \ge$ $\ge f(a + tv)$ for all $t \in H_1 \cap H_2$. Hence the function $f_{a,v} : t \mapsto f(a + tv)$ at 0 has a local maximum on the right, so by the differentiability we obviously get that $f_{a,v}'(0) \le 0$, i.e.,

$$[f'(a)]\, v \le 0,$$

thus, $v \in [f'(a)]^{-1}((-\infty, 0])$. By all these we have just shown that

$$\bigcap_{k \in A} [f_k'(a)]^{-1}((-\infty, 0)) \subseteq [f'(a)]^{-1}((-\infty, 0]),$$

so by the functionals $f_k'(a)$ $(k \in A)$ being sectoroid, we can apply the latest theorem, by which we have that $f'(a)$ is a nonnegative combination of the functionals $\{f_k'(a) : k \in A\}$. Thus, $f'(a)$ is of the form $f'(a) = \sum_{k \in A} \lambda_k f_k'(a)$ where $\lambda_k \geq 0$ $(k \in A)$. Now for the case $k \in B$ by choosing $\lambda_k = 0$ we immediately get the statement. $\Diamond$

References

[1] BROYDEN, C. G.: A simple algebraic proof of Farkas's lemma and related theorems, *Optimization Methods and Software* **8** (3–4) (1998), 185–199.

[2] FARKAS, Gy.: Theorie der einfachen Ungleichungen, *Crelles Journal für die Reine und Angewandte Mathematik* **124** (1902), 1–27.

[3] GALE, D.: The Theory of Linear Economic Models, McGraw-Hill, New York, 1960.

[4] KOMORNIK, V.: A simple proof of Farkas' lemma, *Am. Math. Monthly* December (1998), 949–950.

[5] KUHN, H. W. and A.W. TUCKER: Nonlinear programming, Proceedings of the second Berkeley symposium on mathematical statistics and probability, University of California Press, 1951, 481–492.

[6] PRÉKOPA, A.: On the development of optimization theory, *Amer. Math. Monthly* **87** (1980), 527–542.

[7] RAPCSÁK, T.: Smooth Nonlinear Optimization in R^n, Kluwer Academic Publishers, Dordrecht, 1997.