

ISSN: 2281-1346



Department of Economics and Management

DEM Working Paper Series

Again on the Farkas Theorem and the Tucker Key Theorem Proved Easily

Giorgio Giorgi
(Università di Pavia)

94 (10-14)

Via San Felice, 5
I-27100 Pavia
<http://epmq.unipv.eu/site/home.html>

October 2014

Again on the Farkas Theorem and the Tucker Key Theorem Proved Easily

Giorgio Giorgi (*)

Abstract

The purpose of this paper is twofold; first, to present a simple proof of the Farkas theorem (or Farkas lemma or Farkas-Minkowski lemma), proof performed through a nonlinear theorem of the alternative; second, to present various new proofs of the so-called "Tucker key theorem", and to show that these two results are essentially equivalent.

Key words

Farkas lemma, Farkas theorem of the alternative, Tucker key theorem, Gordan theorem, Stiemke theorem, Motzkin theorem.

1. Introduction

The Tucker key theorem for linear systems can produce many useful results, namely several theorems of the alternative, and is a basic tool in the analysis of some linear economic models. The reader may refer to Koopmans (1951), Howe (1960), Nikaido (1968, 1972), Mangasarian (1969), etc. Even in linear programming and in nonlinear programming many results depend on the so-called Farkas theorem (or lemma) or Farkas-Minkowski theorem (or lemma) or Farkas theorem of the alternative, which in turn can be easily obtained from the Tucker key theorem (see, e. g., Mangasarian (1969)). The original proof of Tucker (1956) of his "key theorem" is purely algebraic and performed by induction. The same proof of Tucker is presented by Mangasarian (1969) and by Kemp and Kimura (1978).

* Department of Economics and Management, Via S. Felice, 5 - 27100 Pavia, (Italy). E-mail: ggiorgi@eco.unipv.it

Nikaido (1968) offers an interesting proof which utilizes the Stiemke theorem of the alternative; the proof of Nikaido is however based on properties of closed convex cones and is a little sophisticated. Singh and Praton (1996) present two proofs of the Tucker key theorem: the first proof relies on results from the theory of convex polytopes, the second proof relies on results of Linear Algebra and is more elementary but more intricate than the first proof. Fujimoto (1976) presents an “autonomous” proof of the key theorem, proof essentially based on properties of an optimization problem.

The aim of the present paper is to develop several proofs of the Tucker key theorem and to give also a proof of the Farkas theorem which does not require to show previously that a convex polyhedral cone is a closed set. This proof is an immediate consequence of a nonlinear theorem of the alternative, which may be considered a nonlinear version of the Farkas theorem.

The paper is organized as follows.

In Section 2 we recall the Tucker key theorem and the Farkas theorem of the alternative. We prove the Farkas theorem by means of a nonlinear theorem of the alternative, considered by Berge and Ghouila-Houri (1965), and here proved in a simple way.

In Section 3 we obtain the Tucker key theorem from the Farkas theorem and the Motzkin theorem, through a general theorem of the alternative obtained by De Giuli, Giorgi and Magnani (1997). This shows that the key theorem and the Farkas theorem are essentially equivalent results, as well as the key theorem and all the known theorems of the alternative for linear systems.

In Section 4 we obtain the key theorem directly from the Stiemke theorem of the alternative, but with an elementary and algebraic proof.

In Section 5 we obtain the Tucker key theorem from another useful theorem, again due to Tucker and concerning skew-symmetric matrices. In Section 6 we present the "direct" proof of the Tucker key theorem, due to Fujimoto (1976) and based on an approach similar to the one of Morishima (1969) for his proof of the Farkas theorem.

Throughout the paper all matrices and vectors are real. The notation $[0]$ stands for the zero (row or column) vector or for the zero matrix. The notations $x \geq y$, $x > y$ (x and y vectors of \mathbb{R}^n) mean, respectively, $x_i \geq y_i$, $\forall i = 1, \dots, n$; $x_i > y_i$, $\forall i = 1, \dots, n$. The notation $x \geq y$ means $x \geq y$, but $x \neq y$. If $y = [0]$, when $x \geq [0]$, $x > [0]$, $x \geq [0]$, we call x , respectively, a nonnegative vector, a positive vector and a semipositive vector. The notations $x \leq y$, $x < y$, $x \leq y$ are obvious.

2. The Tucker key theorem and the Farkas theorem

The Tucker key theorem states, in a slightly more precise version than the usual original version, the following result.

Theorem 1. Given any (m, n) real matrix A and the following two problems:

- (i) Find $x \in \mathbb{R}^n$ such that $Ax = [0]$ and $x \geq [0]$.
- (ii) Find $y \in \mathbb{R}^m$ such that $y^T A \geq [0]$.

Then the problems (i) and (ii) have at least one pair of solutions such that

$$x^\top + y^\top A > [0] \text{ and } x_i(y^\top A)_i = 0 \text{ for all } i.$$

We note at once that the last equality of the thesis is easily obtained: if $x_i = 0$ the equality follows trivially, and if $x_i > 0$, being, from (i), $y^\top Ax = 0$, if $(y^\top A)_i > 0$ we would have a contradiction. So, if $x_i > 0$ it will hold $(y^\top A)_i = 0$.

Obviously, we may have other equivalent reformulations of Theorem 1. For example, we can consider the formulation: the systems

$$Ax \geq [0], \quad A^\top y = [0], \quad y \geq [0]$$

($x \in \mathbb{R}^n, y \in \mathbb{R}^m$) have, respectively, solutions x° and y° such that $Ax^\circ + y^\circ > [0]$.

The well known Farkas theorem of the alternative can be given in the following version.

Theorem 2. Given any (m, n) real matrix A and any vector $b \in \mathbb{R}^m$, then the system

$$S_1 = \{Ax = b, x \geq [0]\}$$

has solutions $x \in \mathbb{R}^n$ if and only if the system

$$S_1^* = \{y^\top A \geq [0], y^\top b < 0\}$$

has no solution $y \in \mathbb{R}^m$.

There are many proofs of this famous result. Perhaps one of the most elementary proofs is the one given by Gale (1960), which, similarly to the proof of Tucker of his “key theorem”, is purely algebraic and performed by induction. Other proofs are usually based on a classical separation property of closed convex sets; however, most of the proofs assume as an obvious fact the closedness of a polyhedral cone. This property is by no means so obvious and it can be proved (see Borwein (1983)) that it is equivalent to the Farkas theorem itself!

Remark 1. The system S_1^* can be equivalently reformulated as

$$S_1^* = \{y^\top A \leq [0], y^\top b > 0\};$$

moreover, S_1^* can be equivalently reformulated as

$$S_1^* = \{A^\top y \geq [0], b^\top y < 0\}.$$

Moreover, the role of S_1 and S_1^* can be interchanged, for example we can consider the formulation

$$\begin{aligned} S_1 &= \{Ax \leq [0], b^\top y > 0\}. \\ S_1^* &= \{A^\top y = b, y \geq [0]\}. \end{aligned}$$

Remark 2. Sometimes the Farkas theorem is not formulated as an alternative theorem, but in the following (equivalent) form (see, e. g. Giannessi (1982), Kemp and Kimura (1978), Martos (1975)).

- A necessary and sufficient condition that the system $S_1 = \{Ax = b, x \geq [0]\}$ has a solution is that the following implication holds

$$y^\top A \geq [0] \implies y^\top b \geq 0.$$

Remark 3. Note that the Farkas theorem is trivial for $b = [0]$, as in this evenience S_1 always has the zero solution and S_1^* is impossible.

We have said that there are many ways to prove the Farkas theorem; roughly speaking, following Broyden (1998) it is possible to distinguish three classes of proofs: algebraic, algorithmic, and geometric. In spite of what promised by some titles, not always the related proofs are short, nor elementary. Without any claim of completeness, we recall the proofs given by Bartl (2008), Broyden (1998), Dax (1997), Good (1959), Nikaido (1968, 1972), Avis and Kaluzny (2004), V. Komornik (1998), Pearl (1967). A good treatment can be found in Bertsekas (1999); finally, there are proofs which use the Tucker's key theorem: see, e. g., Kemp and Kimura (1978), Mangasarian (1969).

We recall also an interesting proof of the Farkas theorem, due to Morishima (1969), proof which is indeed elementary and short and which has never been taken into consideration since now in the mathematical literature, as far as we are aware, perhaps because it appeared in a book on economic analysis.

Here we propose a proof which is an immediate by-product of a general theorem of the alternative for nonlinear systems.

Many theorems of the alternative for nonlinear systems are available in the mathematical literature. These theorems usually hold under various generalized convexity assumptions on the functions involved and some of them are also formulated in an infinite-dimensional topological setting. Curiously, several of these theorems do not recover directly, when applied to the (finite-dimensional) linear case, the Farkas-Minkowski theorem. It is the case, for example, of the results of Bazaraa (1973), of Jeyakumar (1985), of Illés and Kassay (1994) and of Frenk and Kassay (1999). An exception is a theorem presented by Berge and Ghouila-Houri (1965), proved by means of a result, due to Berge, on the intersection of convex sets, result which may be considered an extension of the famous Helly theorem. The elegant proof of Berge and Ghouila-Houri is therefore not quite elementary. Now we prove, with a simple and self-contained procedure, a slight generalization of the theorem of Berge and Ghouila-Houri.

Theorem 3. Let be given the system

$$\begin{cases} f(x) < 0; \\ g_j(x) \leq 0, \quad j = 1, \dots, m; \\ x \in C; \end{cases} \quad (1)$$

where $f : C \rightarrow \mathbb{R}$, $g_j : C \rightarrow \mathbb{R}$ are convex functions on the convex set $C \subset \mathbb{R}^n$. Let us suppose that the following “generalized Slater condition” holds: there exists $x^\circ \in \text{relint}C$ such that

$$\begin{cases} g_j(x^\circ) < 0 \text{ for all } j \text{ such that } g_j \text{ is not linear;} \\ g_j(x^\circ) \leq 0, \text{ for all } j \text{ such that } g_j \text{ is linear.} \end{cases} \quad (2)$$

Then system (1) admits no solution if and only if there exists a vector $y = [y_1, y_2, \dots, y_m] \geq [0]$ such that

$$f(x) + \sum_{j=1}^m y_j g_j(x) \geq 0, \quad \forall x \in C. \quad (3)$$

Before proving Theorem 3 we introduce some notations and some preliminary considerations. Let us denote by F the following set

$$F = \{x \in C : g_j(x) \leq 0, \quad j = 1, \dots, m\}.$$

Some functions g_j may be identically zero on F ; we call these functions “singular functions”, while the others are called “regular functions”. We introduce the following index sets.

$$\begin{aligned} J &= \{1, 2, \dots, m\}; \\ J_s &= \{j \in J : g_j(x) = 0, \quad \forall x \in F\}; \\ J_r &= J \setminus J_s = \{j \in J : g_j(x) < 0 \text{ for some } x \in F\}. \end{aligned}$$

We note then, that if the generalized Slater condition holds, all singular functions g_j must be linear. We recall the following classical separation theorem (see, e. g., Mangasarian (1969), Rockafellar (1970)), which does not require that the set $X \subset \mathbb{R}^n$ (involved in the theorem) is closed.

Lemma 1. Let $X \subset \mathbb{R}^n$ be a nonempty convex set, with $[0] \notin X$. Then, there exists a hyperplane $H = \{x : x \in \mathbb{R}^n, cx = 0\}$, with $c \in \mathbb{R}^n$, $c \neq [0]$, which separates X from the origin of \mathbb{R}^n , i. e. such that $cx \geq 0, \forall x \in X$ and $c\bar{x} > 0$ for some $\bar{x} \in X$.

Proof of Theorem 3. If (1) admits a solution, obviously (3) cannot hold for that solution. This is the trivial part of the proof, which holds without any convexity (or generalized convexity) assumption and without the generalized Slater condition.

Now, let us assume that (1) does not hold. With $u = [u_0, u_1, \dots, u_m] \in \mathbb{R}^{m+1}$, let us define the following set

$$U = \{u : \exists x \in C \text{ such that } u_0 > f(x), \quad u_j \geq g_j(x) \text{ if } j \in J_r, \quad u_j = g_j(x), \text{ if } j \in J_s\}.$$

Clearly, U is convex and does not contain the origin of \mathbb{R}^{m+1} . Therefore, thanks to Lemma 1, there exists a separating hyperplane, defined by the nonzero vector $[y_0, y_1, \dots, y_m]$, such that

$$\sum_{j=0}^m y_j u_j \geq 0, \quad \forall u \in U \quad (4)$$

and

$$\sum_{j=1}^m y_j \bar{u}_j > 0, \text{ for some } \bar{u}_j \in U. \quad (5)$$

We perform the remaining proof in three steps:

- (I) First we prove that $y_0 \geq 0$ and $y_j \geq 0$ for all $j \in J_r$.
- (II) Secondly we establish that (4) and (5) hold for $u = (f(x), g_1(x), \dots, g_m(x))$ if $x \in C$.
- (III) Then we prove that $y_0 > 0$.

Proof of (I). We show that $y_j \geq 0$ for all $j \in \{0\} \cup J_r$. Let us assume that $y_0 < 0$. Let us take an arbitrary vector $(u_0, u_1, \dots, u_m) \in U$. By definition $(u_0 + \lambda, u_1, \dots, u_m) \in U$ for all $\lambda \geq 0$. Hence by (4) one has

$$\lambda y_0 + \sum_{j=0}^m y_j u_j \geq 0 \text{ for all } \lambda \geq 0.$$

For sufficiently large λ the left hand side of the last inequality is negative, which is a contradiction. Therefore it holds $y_0 \geq 0$. The proof of the nonnegativity of all other y_j , $j \in J_r$, is similar.

Proof of (II). Secondly, we establish that

$$y_0 f(x) + \sum_{j=1}^m y_j g_j(x) \geq 0 \text{ for all } x \in C. \quad (6)$$

This follows from the remark that for all $x \in C$ and for all $\lambda \geq 0$ one has $u = (f(x) + \lambda, g_1(x), \dots, g_m(x)) \in U$, thus

$$y_0(f(x) + \lambda) + \sum_{j=1}^m y_j g_j(x) \geq 0 \text{ for all } x \in C.$$

Taking the limit as $\lambda \rightarrow 0$ the claim follows.

Proof of (III). Thirdly we show that $y_0 > 0$. The proof is by contradiction. We already know that $y_0 \geq 0$. Let us assume that $y_0 = 0$. Hence, from (6) we have

$$\sum_{j \in J_r} y_j g_j(x) + \sum_{j \in J_s} y_j g_j(x) = \sum_{j=1}^m y_j g_j(x) \geq 0 \text{ for all } x \in C.$$

Taking a point $x^* \in \text{relint}C$, such that

$$\begin{aligned} g_j(x^*) &< 0, \quad \forall j \in J_r \\ g_j(x^*) &= 0, \quad \forall j \in J_s \end{aligned}$$

(this point surely exists, thanks to (2)) one has

$$\sum_{j \in J_r} y_j g_j(x^*) \geq 0.$$

Since $y_j \geq 0$ and $g_j(x^*) < 0$ for all $j \in J_r$, this implies $y_j = 0$ for all $j \in J_r$. Therefore it holds

$$\sum_{j \in J_s} y_j g_j(x) \geq 0 \text{ for all } x \in C. \quad (7)$$

Now, from (5), with $\bar{x} \in C$ such that $\bar{u}_j = g_j(\bar{x})$ for $j \in J_s$, we have

$$\sum_{j \in J_s} y_j g_j(\bar{x}) > 0. \quad (8)$$

Because $x^* \in \text{relint}C$, there exist a vector $\tilde{x} \in C$ and $0 < \lambda < 1$ such that $x^* = \lambda\bar{x} + (1-\lambda)\tilde{x}$. Taking into account that it holds $g_j(x^*) = 0$ for $j \in J_s$ and that the singular functions are linear, one gets

$$\begin{aligned} 0 &= \sum_{j \in J_r} y_j g_j(x^*) = \sum_{j \in J_s} y_j g_j(\lambda\bar{x} + (1-\lambda)\tilde{x}) = \\ &= \lambda \sum_{j \in J_s} y_j g_j(\bar{x}) + (1-\lambda) \sum_{j \in J_s} y_j g_j(\tilde{x}) > (1-\lambda) \sum_{j \in J_s} y_j g_j(\tilde{x}). \end{aligned}$$

The last inequality follows from (8). Since $(1-\lambda) > 0$ we obtain the inequality

$$\sum_{j \in J_s} y_j g_j(\tilde{x}) < 0$$

which contradicts (7). Hence we have proved that $y_0 > 0$.

At this point we have (6), with $y_0 > 0$ and $y_j \geq 0$ for all $j \in J_r$. Dividing by $y_0 > 0$ in (6) and by defining $y_j \equiv (y_j/y_0)$ for all $j \in J$ we obtain the thesis. \square

Remark 4. Of course the multipliers of all singular constraints can always be chosen strictly positive.

Remark 5. The classical Farkas-Minkowski theorem, where, e. g., S_1 and S_1^* have the form

$$\begin{aligned} S_1 : \quad & Ax = b, \quad x \geq [0] \\ S_1^* : \quad & A^\top u \geq [0], \quad b^\top u < 0, \end{aligned}$$

can be simply deduced from Theorem 3. It is quite immediate to verify that S_1 and S_1^* cannot have both a solution. It remains to prove that if S_1^* has no solutions, then S_1

admits a solution. Let us write S_1^* in the form $f(u) \equiv b^\top u < 0, -A^\top u \leq [0]$. If S_1^* admits no solutions, from Theorem 3 we have that there exists $x \in \mathbb{R}_+^n$ such that

$$b^\top u - x^\top A^\top u = u^\top (b - Ax) \geq [0],$$

for each $u \in \mathbb{R}^m$, and therefore it holds $b - Ax = [0], x \geq [0]$. \square

Remark 6. In Giorgi (2002) it is proved the following generalization of Theorem 3, by use of the Fan-Glicksberg-Hoffman theorem of the alternative (see, e. g., Mangasarian (1969)), which may be considered a nonlinear version of the Gordan theorem of the alternative and which can be proved in a quite elementary way.

Theorem 4. Let $C \subset \mathbb{R}^n$ be a nonempty convex set, $f : C \longrightarrow \mathbb{R}^p, g : C \longrightarrow \mathbb{R}^q$ be convex functions and $h : \mathbb{R}^n \longrightarrow \mathbb{R}^r$ be a linear function. Let us assume that there exists $\bar{x} \in \text{relint}C$ such that $g_j(\bar{x}) < 0, j = 1, \dots, q$, and $h_i(\bar{x}) \leq 0, i = 1, \dots, r$. Then the system

$$\begin{cases} f_k(x) < 0, & k = 1, \dots, p; \\ g_j(x) \leq 0, & j = 1, \dots, q; \\ h_i(x) \leq 0, & i = 1, \dots, r; \end{cases}$$

admits no solutions if and only if there exists a vector $(u, v, w)^\top \in \mathbb{R}_+^p \times \mathbb{R}_+^q \times \mathbb{R}_+^r, u \neq [0]$, such that

$$u^\top f(x) + v^\top g(x) + w^\top h(x) \geq 0, \forall x \in C.$$

From this theorem it is quite easy to deduce the Motzkin theorem of the alternative (see the next Theorem 6). Another generalized Motzkin theorem has been obtained by Jeyakumar (1985), under a regularity assumption, due to Karlin, which is equivalent to Slater condition.

Remark 7. Other nonlinear theorems of the alternative, useful to deduce directly the Farkas-Minkowski theorem, are due to Cambini (1986) and Giannessi (1980, 1984). In particular, the following result is a particular case of a more general theorem proved by the said authors.

Theorem 5. Let $C \subset \mathbb{R}^n$ be a nonempty convex set and let $f : C \longrightarrow \mathbb{R}^p, g : C \longrightarrow \mathbb{R}^q$ be convex functions. Then:

(i) If the system

$$\begin{cases} f_k(x) < 0, & k = 1, \dots, p; \\ g_j(x) \leq 0, & j = 1, \dots, q; \end{cases} \quad (9)$$

admits no solutions, then there exist vectors $u^\top \in \mathbb{R}_+^p, v^\top \in \mathbb{R}_+^q$, with $(u, v)^\top \neq [0]$, such that

$$u^\top f(x) + v^\top g(x) \geq 0, \forall x \in C. \quad (10)$$

(ii) If (10) holds with $u^\top \in \mathbb{R}_+^p, v^\top \in \mathbb{R}_+^q, (u, v)^\top \neq [0]$ and, moreover, if it holds

$$\{x \in C, f_k(x) < 0, k = 1, \dots, p; g_j(x) \leq 0, j = 1, \dots, q; v^\top g(x) = 0\} = \emptyset,$$

whenever $u = [0]$, then system (9) is impossible.

From this theorem Giannessi (1980) obtains in a simple way a non-homogeneous version of the Farkas-Minkowski theorem, due to Duffin (see Mangasarian (1969)). From the Duffin theorem at once the Farkas-Minkowski theorem follows.

3. The Tucker key theorem from the Farkas theorem

We first obtain the *Motzkin theorem of the alternative* from the Farkas theorem. We recall the Motzkin theorem of the alternative.

Theorem 6. Given (real) matrices A , B and H of appropriate dimensions, exactly one of the following two systems has a solution:

- (i) $Ax < [0]$, $Bx \leq [0]$, $Hx = [0]$.
- (ii) $u^\top A + v^\top B + w^\top H = [0]$, $u \geq [0]$, $v \geq [0]$.

Proof.

a) It is easy to show that both (i) and (ii) cannot have a solution. Suppose $u^\top A + v^\top B + w^\top H = [0]$ for some $u \geq [0]$, $v \geq [0]$, w unrestricted in sign. Then, for every vector x we have $u^\top Ax + v^\top Bx + w^\top Hx = 0$. If $Bx \leq [0]$, then $v^\top Bx \leq 0$ and if $Hx = [0]$, then $w^\top Hx = 0$. Thus $u^\top Ax \geq 0$. Since u is semipositive, $Ax < [0]$ cannot hold.

b) Suppose now that (i) has no solution. Then the system

$$\begin{cases} Ax + e\theta \leq [0], \theta > 0; \\ Bx \leq [0]; \\ Hx \leq [0]; \\ -Hx \leq [0]; \end{cases}$$

has no solution (the vector e is the summing vector, i. e. $e = [1, 1, \dots, 1]^\top$). This system can be rewritten in the form

$$\begin{bmatrix} A & e \\ B & [0] \\ H & [0] \\ -H & [0] \end{bmatrix} \begin{pmatrix} x \\ \theta \end{pmatrix} \leq [0], \quad (0, \dots, 0, 1) \begin{pmatrix} x \\ \theta \end{pmatrix} > 0.$$

From the Farkas theorem, there exists a vector $(u, v, w^1, w^2) \geq [0]$ such that

$$\begin{bmatrix} A & e \\ B & [0] \\ H & [0] \\ -H & [0] \end{bmatrix}^\top \begin{pmatrix} u \\ v \\ w^1 \\ w^2 \end{pmatrix} = \begin{pmatrix} 0 \\ \cdot \\ 0 \\ 1 \end{pmatrix}.$$

This can be rewritten as $u^\top A + v^\top B + (w^1 - w^2)^\top H = [0]$, $u^\top e = 1$. Letting $w^\top = (w^1 - w^2)^\top$, we have completed the proof of b). \square

Remark 8. Obviously also the Motzkin theorem of the alternative can be equivalently reformulated in other forms, e. g., as an alternative between

$$S_2 = \{Ax > [0], \quad Bx \geq [0], \quad Hx = [0]\}$$

and

$$S_2^* = \{u^\top A + v^\top B + w^\top H = [0], \quad u \geq [0], \quad v \geq [0]\}.$$

By means of the Motzkin theorem of the alternative and with algebraic operations only, De Giuli, Giorgi and Magnani (1997) obtained a general theorem of the alternative for linear systems which contains *all* the known theorems of the alternative for linear systems and many other formulations: on the whole 225 theorems of the alternative!

Let the real matrix A of order (m, n) , and the column vectors $b \in \mathbb{R}^m$ and $x \in \mathbb{R}^n$ be partitioned in the following forms

$$A = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1q} \\ A_{21} & A_{22} & \dots & A_{2q} \\ \dots & \dots & \dots & \dots \\ A_{p1} & A_{p2} & \dots & A_{pq} \end{bmatrix}, \quad b = \begin{bmatrix} b^1 \\ b^2 \\ \vdots \\ b^p \end{bmatrix}, \quad x = \begin{bmatrix} x^1 \\ x^2 \\ \vdots \\ x^q \end{bmatrix},$$

with A_{ij} of order (m_i, n_j) , $b^i \in \mathbb{R}^{m_i}$, $x^j \in \mathbb{R}^{n_j}$, $p > 3, q > 3$. Define the submatrices

$$A_i = [A_{i1}; A_{i2}; \dots; A_{iq}], \quad i \in \{1, 2, \dots, p\}$$

$$A^j = \begin{bmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ A_{pj} \end{bmatrix}, \quad j \in \{1, 2, \dots, q\}$$

and set up the system

$$S_3 = \begin{cases} A_1 x = b^1 \\ A_2 x \leq b^2 \\ A_3 x < b^3 \\ A_i x \leq b^i, \quad i \in \{4, 5, \dots, p\} \\ x^1 \text{ sign unrestricted} \\ x^2 \geq [0] \\ x^3 > [0] \\ x^j \geq [0], \quad j \in \{4, 5, \dots, q\}. \end{cases}$$

The following general theorem of the alternative is valid for system S_3 (see De Giuli, Giorgi and Magnani (1997)).

Theorem 7. System S_3 admits a solution x if and only if no vector

$$y^\top = [(y^1)^\top; (y^2)^\top; (y^3)^\top; (y^4)^\top; \dots; (y^p)^\top],$$

$(y^i)^\top \in \mathbb{R}^{m_i}$, $i \in \{1, 2, \dots, p\}$, solves the system

$$S_3^* = \left\{ \begin{array}{ll} y^\top A^1 = [0] & \\ y^\top A^2 \geq [0] & \\ y^\top A^3 \geq [0] & y^\top A^3 \geq [0] \\ y^\top A^j \geq [0], \quad j \in \{4, 5, \dots, q\} & y^\top A^j > [0] \text{ for some } j \geq 4 \\ (y^1)^\top \text{ sign unrestricted} & \\ (y^2)^\top \geq [0] & \\ (y^3)^\top \geq [0] & (y^3)^\top \geq [0] \\ (y^i)^\top \geq [0], \quad i \in \{4, 5, \dots, p\} & (y^i)^\top > [0] \text{ for some } i \geq 4 \\ y^\top b \leq 0 & y^\top b < 0 \\ \text{moreover, at least one of the relations} & \\ \text{of the second column holds.} & \end{array} \right.$$

The effort to prove Theorem 7 pays us, because, as we have already mentioned, the said theorem can generate 225 theorems of the alternative, among which all the "classical" theorems of the alternative for linear systems, already known in the literature (see, e. g., Mangasarian (1969), Giannessi (1982, 2005)). For other considerations the reader is referred to the paper of De Giuli, Giorgi and Magnani (1997).

From Theorem 7 it is possible to deduce at once the Tucker key theorem: we have to prove that the systems

$$Ax \geq [0], \quad A^\top y = [0], \quad y \geq [0]$$

($x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$) have, respectively, solutions x° and y° such that $Ax^\circ + y^\circ > [0]$. In other words, we have to prove that the system

$$\left\{ \begin{array}{l} A^\top y = [0] \\ -Ax \leq [0] \\ -Ax - Iy > [0] \\ y \geq [0] \end{array} \right.$$

admits a solution. On the ground of Theorem 7, it is sufficient to prove that its "dual" system

$$\left\{ \begin{array}{l} A^\top u^2 + A^\top u^3 = [0] \\ Au^1 - u^3 \geq [0] \\ u^1 \in \mathbb{R}^n, u^2 \geq [0], u^3 \geq [0] \end{array} \right.$$

admits no solution.

Let us absurdly suppose that this last system admits a solution. We obtain

$$0 = (u^1)^\top A^\top (u^2 + u^3) = (u^2 + u^3)^\top Au^1 \geq (u^2 + u^3)^\top u^3 \geq (u^3)^\top u^3 > 0,$$

which is a contradiction. So, the Tucker key theorem is proved. \square

We have mentioned that Tucker's original proof is by induction and quite elementary: more precisely, Tucker first proves the following result: the system

$$Ax = [0], \quad x \geq [0], \quad A^\top y \geq [0] \quad (11)$$

has solutions x and y such that

$$(x + A^\top y)_1 > 0.$$

We note that this preliminary result can be easily obtained from the Farkas theorem, which shows to be once more a "river from which many streams flow". Indeed, if there exists a solution of (11) with $x_1 > 0$, there is nothing to prove. If there does not exist a solution with $x_1 > 0$, then writing $x_1 = (e^1)^\top x$, where $(e^1)^\top = [1, 0, \dots, 0]$, the system

$$Ax = [0], \quad x \geq [0], \quad (-e^1)^\top x < 0$$

does not have a solution; then Farkas theorem states that the system $A^\top z \leq -e^1$ has a solution. Hence, with $y = -z$, one has a solution of (11) such that the first coordinate of $(x + A^\top y)$ is positive. From this preliminary result the Tucker key theorem follows easily: let us define

$$x^\circ = \sum_{j=1}^n x^j \quad \text{and} \quad y^\circ = \sum_{j=1}^n y^j,$$

where $x^j \in \mathbb{R}^n$ and $y^j \in \mathbb{R}^m$ are solutions of (1), their existence following from the preliminary result. Then

$$x_i^\circ + (A^\top y^\circ)_i \geq (x^i + A^\top y)_i > 0, \quad i = 1, \dots, n.$$

4. The Tucker key theorem from the Stiemke theorem

We recall the Stiemke theorem of the alternative.

Theorem 8. (Stiemke theorem of the alternative). Given any matrix A of order (m, n) , the system

$$Ax \geq [0]$$

admits a solution if and only if the system

$$A^\top y = [0], \quad y > [0]$$

does not admit solution.

The Stiemke theorem can be obtained by the Tucker key theorem. Here we shall prove the converse (the fact that the two theorems are in fact equivalent comes also from the results of the previous section).

Lemma 2. The system $Ax \geq [0]$ can always be decomposed as follows

$$\{x \mid Ax \geq [0]\} = \left\{ x \mid \begin{array}{l} A_{(1)}x = [0] \\ A_{(2)}x \geq [0] \end{array} \right\}, \quad (12)$$

$$\left\{ x \mid \begin{array}{l} A_{(1)}x = [0] \\ A_{(2)}x > [0] \end{array} \right\} \neq \emptyset, \quad (13)$$

$$\{x \mid A_{(1)}x \geq [0]\} = \emptyset, \quad (14)$$

where one of the submatrices $A_{(1)}$ or $A_{(2)}$ can be empty.

Proof. We build up the sets

$$Q_i = \{x \mid Ax \geq [0], A_i x > 0\}, i = 1, \dots, m,$$

and decompose the matrix A into the two submatrices

$$\begin{aligned} A_{(1)} &= (A_i)_{i \in I_1}, & I_1 &= \{i \mid Q_i = \emptyset\} \\ A_{(2)} &= (A_i)_{i \in I_2}, & I_2 &= \{i \mid Q_i \neq \emptyset\}. \end{aligned}$$

We note that, owing to this construction, relation (12) is satisfied, as from $Ax \geq [0]$ it follows $A_i x = 0$ for all $i \in I_1$, that is $A_{(1)}x = [0]$. Then, relation (13) follows, as, owing to the above construction, there exist vectors $q^i \in Q_i$, $i \in I_2$, and for $q = \sum q^i$ we have $A_{(1)}q = [0]$, $A_{(2)}q > [0]$.

Therefore q is an element of the set considered in (13). The property (14) will be proved in an indirect way. If it holds $A_{(1)}z \geq [0]$ for a given z , then it will follow, for $x = \lambda q + z$, being q the vector defined above and $\lambda > 0$,

$$\begin{aligned} A_{(1)}x &= A_{(1)}z \geq [0] \\ A_{(2)}x &= \lambda A_{(2)}q + A_{(2)}z \geq [0] \end{aligned}$$

where the second row is obtained thanks to the fact that it holds $A_{(2)}q > [0]$ and with λ sufficiently large. But then x contradicts relation (12), i. e. x is in the set described by the left-hand side of (12), but not in the set described in the right-hand side of (12). \square

Proof of the Tucker key theorem. Let us decompose the system $Ax \geq [0]$ on the ground of Lemma 2. From (13) and (14) we get

$$\exists x^\circ \in \mathbb{R}^n : A_{(1)}x^\circ = [0], \quad A_{(2)}x^\circ > [0] \quad (15)$$

$$\nexists x \in \mathbb{R}^n : A_{(1)}x \geq [0]. \quad (16)$$

From the Stiemke theorem and from (16) it follows

$$\exists y^1 > [0] : A_{(1)}^\top y^1 = [0].$$

Then we obtain, taking relation (15) into account,

$$\exists x^\circ, y^\circ = \begin{bmatrix} y^1 \\ [0] \end{bmatrix} : Ax^\circ + y^\circ = \begin{bmatrix} [0] + y^1 \\ A_{(2)}x^\circ + [0] \end{bmatrix} > [0]$$

with $Ax^\circ \geq [0]$, $A^\top y^\circ = [0]$, $y^\circ \geq [0]$, i. e. the Tucker key theorem. \square

The Stiemke theorem of the alternative is in fact equivalent to the *Gordan theorem of the alternative* (Gordan (1873)), which is perhaps the first theorem of the alternative for linear systems published on a mathematical journal.

Theorem 9. (Gordan theorem of the alternative). Given any matrix A of order (m, n) , the system

$$Ax > [0]$$

admits a solution $x \in \mathbb{R}^n$ if and only if the system

$$A^\top y = [0], \quad y \geq [0]$$

does not admit a solution $y \in \mathbb{R}^m$.

The said equivalence was noted by Antosiewicz (1955) for more general systems and can be proved easily as follows. By means of the matrix A , let us construct the following linear spaces (complementary to each other):

$$\begin{aligned} L &= \{v \in \mathbb{R}^m : v = Ax, x \in \mathbb{R}^n\} \\ L^\perp &= \{y \in \mathbb{R}^m : A^\top y = [0]\}. \end{aligned} \quad (17)$$

We can then rewrite the Gordan theorem in the “alternative form”:

$$\begin{aligned} (I) \quad & \exists v \in L, v > [0]; \\ (II) \quad & \exists y \in L^\perp, y \geq [0]. \end{aligned} \quad (18)$$

Being $L^{\perp\perp} = L$, we can dualize the alternative relations (18), by exchanging the order relations on the respective vectors:

$$\begin{aligned} (I') \quad & \exists v \in L, v \geq [0]; \\ (II') \quad & \exists y \in L^\perp, y > [0]. \end{aligned} \quad (19)$$

If we express (19) by means of (17), we obtain a dual version of the Gordan theorem of the alternative, i. e. the Stiemke theorem of the alternative:

$$(I'') \quad \exists x \in \mathbb{R}^n : Ax \geq [0];$$

or

$$(II'') \quad y \in \mathbb{R}^m : A^\top y = [0], y > [0],$$

but never both.

The proof of the Gordan theorem of the alternative by means of a separation theorem, has the advantage, with respect to the similar proof of the Farkas theorem, to require a weak separation theorem, where no “closedness property” is required. See, e. g., Bazaraa, Sherali and Shetty (1993).

Also the Stiemke theorem of the alternative can be proved easily in a direct way:

a) $(I'') \implies \text{not}(II'')$: clearly both (I'') and (II'') cannot be true, for then we must have both $y^\top Ax = 0$ (as $y^\top A = [0]$) and $y^\top Ax > 0$ (as $y > [0]$ and $Ax \geq [0]$).

b) $\text{not}(I'') \implies (II'')$: let $\Delta = \left\{ z \in \mathbb{R}^m : z \geq [0] \text{ and } \sum_{j=1}^n z_j = 1 \right\}$ be the unit simplex in \mathbb{R}^m . In geometric terms, (I'') asserts that the span M of the columns $\{A^1, \dots, A^n\}$ intersects the nonnegative orthant \mathbb{R}_+^m at a nonzero point, namely Ax . Since M is a linear subspace, we may rescale x so that Ax belongs to $M \cap \Delta$. Thus the negation of (I'') is equivalent to the disjointness of M and Δ . So, assume that (I'') fails. Then, since Δ is compact and convex and M is closed and convex, there is a hyperplane strongly separating Δ and M . That is, there is some nonzero $y \in \mathbb{R}^m$ and some $\epsilon > 0$ satisfying

$$yp + \epsilon < yz \text{ for all } p \in M, z \in \Delta.$$

Since M is a linear subspace, we must have $yp = 0$ for all $p \in M$. Consequently, $yz > \epsilon > 0$ for all $z \in \Delta$. Since the j -th unit coordinate vector e^j belongs to Δ , we see that $y_j = y^\top e^j > 0$, that is $y > [0]$. Since each $A^i \in M$, we have that $y^\top A^i = 0$, i. e. $y^\top A = [0]$. This completes the proof. \square

5. The Tucker key theorem from another Tucker theorem

In the same paper where the key theorem was proved, Tucker presented another useful result, we call the *Tucker existence lemma* for skew-symmetric matrices. This lemma is important to derive duality results for linear programming problems, but can be used also to derive the Tucker key theorem.

Lemma 3. (Tucker existence lemma). Let L be a (square) skew-symmetric (real) matrix (i. e. $L^\top = -L$). Then the system

$$Lx \geq [0], \quad x \geq [0] \tag{20}$$

has a solution \bar{x} for which

$$L\bar{x} + \bar{x} > [0]. \tag{21}$$

Proof. This lemma can be proved directly from the "all-purpose" Theorem 7 of Section 3, by noting that the following systems are in alternative:

- (i) $Ax \geq [0], \quad Bx > [0], \quad x \geq [0];$
- (ii) $A^\top u + B^\top v \leq [0], \quad u \geq [0], v \geq [0].$

Then, in order to prove that the system

$$Lx \geq [0], \quad (L + I)x > [0], \quad x \geq [0]$$

admits a solution, it is sufficient to prove that its "dual" system

$$u \geq [0], v \geq [0], L^\top u + (L^\top + I)v \leq [0]$$

does not admit solutions. Being $L^\top = -L$, the last inequality can be rewritten as $v \leq L(u + v)$. Let us absurdly suppose that such a solution exists. We obtain

$$0 < v^\top v \leq (u + v)^\top v \leq (u + v)^\top L(u + v) = 0,$$

which is obviously absurd. \square

It is, however, possible to prove the Tucker existence lemma directly from the Farkas theorem.

Let L be an arbitrary skew-symmetric matrix, of order n , let e^i be the i -th unit coordinate vector and I the identity matrix of order n . By Farkas theorem, either the system

$$\{-L^\top x \geq [0], \quad Ix \geq [0], \quad (-e^i)^\top x < 0\} \quad (22)$$

has a solution, or the system

$$\{Lv - Iz = e^i, \quad v \geq [0], z \geq [0]\}$$

has a solution y , with $y^\top = (v^\top; z^\top)$, but never both. In the first case, taking into account the equality $L^\top = -L$, we have a solution x^i of (22) for which $Lx^i \geq [0]$, $x^i \geq [0]$, $x^i_i > 0$. In the second case, we have a vector v^i for which $Lv^i \geq e^i$, $v^i \geq [0]$. Therefore, in either case there exists a vector x^i for which $Lx^i + x^i \geq [0]$ and the i -th component of the vector on the left-hand side is positive. Thus, the vector

$$x = \sum_{i=1}^n x^i$$

meets (20) and (21). \square

Theorem 10. Lemma 3 implies the Tucker key theorem.

Proof. Let A be an arbitrary matrix. Let us rewrite the system

$$Ax \geq [0], \quad A^\top y = [0], \quad y \geq [0]$$

in an equivalent matrix form:

$$\begin{bmatrix} [0] & A & -A \\ -A^\top & [0] & [0] \\ A^\top & [0] & [0] \end{bmatrix} \begin{bmatrix} y \\ x^1 \\ x^2 \end{bmatrix} \geq [0], \quad y \geq [0], x^1 \geq [0], x^2 \geq [0]. \quad (23)$$

The matrix appearing in (23) is skew-symmetric. By applying Lemma 3, there exists a solution y, x^1, x^2 of (23) such that

$$\begin{bmatrix} [0] & A & -A \\ -A^\top & [0] & [0] \\ A^\top & [0] & [0] \end{bmatrix} \begin{bmatrix} y \\ x^1 \\ x^2 \end{bmatrix} + \begin{bmatrix} y \\ x^1 \\ x^2 \end{bmatrix} > [0].$$

From this we have $A(x^1 - x^2) + y > [0]$ and setting $x = x^1 - x^2$ we get the thesis of the Tucker key theorem. \square

6. A direct simple proof of the Tucker key theorem

In this section we present, for the reader's convenience, the proof of Fujimoto (1976) of the Tucker key theorem, proof based on a minimization problem and similar to the proof of Morishima (1969) of the Farkas theorem.

Proof of Theorem 1. (Fujimoto (1976)). If (i) of Theorem 1 has a solution which is strictly positive, $x > [0]$, then put $y = [0]$ and we have the desired result. Next, suppose that (i) of Theorem 1 has no solution such that $x > [0]$. Then, there is a vector $\bar{x} = [\bar{x}_1, \dots, \bar{x}_n]^T$ which has the minimum number of zero element(s) among the solutions of (i). Denote this minimum number as r and, without loss of generality, suppose that

$$\bar{x}_j = 0, \text{ for } 1 \leq j \leq r$$

and

$$\bar{x}_j > 0, \text{ for } r + 1 \leq j \leq n.$$

Of course, r may be equal to n , in which case there exists the zero solution only. Let be

$$M = x^T A^T A x. \quad (24)$$

Now, consider the following minimization problem

$$(P) \quad \begin{cases} \text{Minimize } M \text{ subject to } x_j \geq 0 \text{ for each } j \text{ and} \\ \sum_{j=1}^r x_j = 1. \end{cases} \quad (25)$$

Note that the variables are the elements x_j and that the summation (25) in the constraint ranges from $j = 1$ to $j = r$. Since the above problem is quadratic, it has a solution vector x^* . It can be shown that the minimum value of M is positive, $M^* > 0$. Because, if it would be zero, we get $Ax^* = [0]$ and we can form a new vector $x^\circ = \bar{x} + x^*$, which satisfies $Ax^\circ = [0]$, $x^\circ \geq [0]$ and has a less number of zero elements than \bar{x} , thus a contradiction. Now we apply the Lagrange multiplier theorem, taking into account that our problem has nonnegative variables. We form the Lagrangian function

$$L = M - \lambda \left(\sum_{j=1}^r x_j - 1 \right).$$

At the solution point x^* , there is the corresponding multiplier value λ^* and we have

$$2x^{*\top} A^T A - \lambda^* e^\circ \geq [0] \quad (26)$$

$$2x^{*\top} A^T A x^* - \lambda^* e^\circ x^* = 0, \quad (27)$$

where $e^\circ = [1, \dots, 1, 0, \dots, 0]$, i. e. the first r elements are the unity, the others are the zero element. Using (24) and (25), the above equation (27) leads to

$$2M^* - \lambda^* = 0.$$

Thus, we find $\lambda^* > 0$. Then, by putting $\bar{y}^\top = 2x^{*\top}A^\top$, it follows from (26) that $\bar{y}^\top A \geq [0]$, where the first r elements are positive. Therefore, we obtain $\bar{x}^\top + \bar{y}^\top A > [0]$. Moreover, since $A\bar{x} = [0]$, the vector $x^* + \bar{x}$ is again a solution *whose last $n - r$ elements are strictly positive*. Supposing x^* is already such a solution, then from (26) and (27), $(\bar{y}^\top A)_i = 0$ for $r + 1 \leq i \leq n$. Thus, $\bar{x}_i(\bar{y}^\top A)_i = 0$ for all i . \square

References

- H. A. Antosiewicz (1955), A theorem on alternatives for pair of matrices, *Pacific Journal of Mathematics*, 5, 641-642.
- D. Avis, B. Kaluzny (2004), Solving inequalities and proving Farkas's lemma made easy, *American Math. Monthly*, 111, 152-157.
- D. Bartl (2008), A short algebraic proof of the Farkas lemma, *SIAM J. Optim.*, 19, 234-239.
- M. S. Bazaraa (1973), A theorem of the alternative with application to convex programming: optimality, duality and stability, *J. Math. Anal. Appl.*, 41, 701-715.
- M. S. Bazaraa, H. D. Sherali and C. M. Shetty (1993), *Nonlinear Programming*, John Wiley & Sons, New York.
- C. Berge, A. Ghouila-Houri (1965), *Programming, Games and Transportation Networks*, Methuen, London.
- D. P. Bertsekas (1999), *Nonlinear Programming (Second Edition)*, Athena Scientific, Belmont, Mass.
- J. M. Borwein (1983), A note on the Farkas lemma, *Utilitas Mathematica*, 24, 235-241.
- C. G. Broyden (1998), A simple algebraic proof of Farkas's lemma and related theorems, *Optim. Methods Softw.*, 8, 185-199.
- A. Cambini (1986), Nonlinear separation theorems, duality and optimality conditions; in R. Conti, E. De Giorgi, F. Giannessi (Eds.), *Optimization and Related Fields*, Springer Verlag, Berlin, 57-93.
- A. Dax (1997), An elementary proof of Farkas' lemma, *SIAM Rev.*, 39, 503-507.
- M. E. De Giuli, G. Giorgi and U. Magnani (1997), A general theorem of the alternative: how to get its special cases quickly, *Pure Mathematics and Applications*, 8, 215-232.
- J. B. G. Frenk, G. Kassay (1999), On classes of generalized convex functions, Gordan-Farkas type theorems and Lagrangian duality, *J. Optim. Theory Appl.*, 102, 315-343.
- T. Fujimoto (1976), A simple proof of the Tucker theorem, *The Journal of Economic Studies*, Toyama University, 21, 269-272.
- D. Gale (1960), *The Theory of Linear Economic Models*, McGraw-Hill, New York.
- F. Giannessi (1980), Theorems of the alternative, quadratic programs and complementarity problems: in R. W. Cottle, F. Giannessi, J. L. Lions (Eds.), *Variational Inequalities and Complementarity Problems*, J. Wiley & Sons, New York, 151-186
- F. Giannessi (1982), *Metodi Matematici della Programmazione. Problemi Lineari e Non Lineari*, Pitagora Editrice, Bologna.
- F. Giannessi (1984), Theorems of the alternative and optimality conditions, *J. Optim. Theory Appl.*, 43, 331-365. Errata Corrige in *J. Optim. Theory Appl.*, 44, 1984, 363-364.
- F. Giannessi (2005), *Constrained Optimization and Image Space Analysis. Volume 1. Separation of Sets and Optimality Conditions*, Springer New York.

- G. Giorgi (2002), Osservazioni sui teoremi dell'alternativa non lineari implicanti relazioni di uguaglianza e vincolo insiemistico; in G. Crespi and others (Eds.), *Optimality in Economics, Finance and Industry*, Datanova, Milan, 171-183.
- R. A. Good (1959), Systems of linear relations, *SIAM Rev.*, 1, 1-31.
- P. Gordan (1873), Ueber die Auflösung linearer Gleichungen mit reellen Coefficienten, *Math. Annalen*, 6, 23-28.
- C. W. Howe (1960), An alternative proof of the existence of general equilibrium in a von Neumann model, *Econometrica*, 28, 635-637.
- T. Illés, G. Kassay (1994), Farkas type theorems for generalized convexities, *Pure Mathematics and Applications*, 5, 225-239.
- V. Jeyakumar (1985), Convexlike alternative theorems and mathematical programming, *Optimization*, 16, 643-652.
- M. C. Kemp, Y. Kimura (1978), *Introduction to Mathematical Economics*, Springer Verlag, Berlin.
- V. Komornick (1998), A simple proof of Farkas' lemma, *American Math. Monthly*, 105, 949-950.
- T. C. Koopmans (Ed.) (1951), *Activity Analysis of Production and Allocation*, John Wiley & Sons, New York.
- O. L. Mangasarian (1969), *Nonlinear Programming*, McGraw-Hill, New York.
- B. Martos (1975), *Nonlinear programming. Theory and Methods*, North Holland, Amsterdam.
- M. Morishima (1969), *Theory of Economic Growth*, Clarendon Press, Oxford.
- H. Nikaido (1968), *Convex Structures and Economic Theory*, Academic Press, New York.
- H. Nikaido (1972), *Introduction to Sets and Mappings in Modern Economics*, North Holland, Amsterdam.
- M. H. Pearl (1967), A matrix proof of the Farkas's theorem, *Quarterly Journal of Math.*, 18, 193-197.
- R. T. Rockafellar (1970), *Convex Analysis*, Princeton University Press, Princeton.
- C. Singh, I. Praton, (1996), Some variations, extensions and generalizations of Tucker's theorem, *Journal of Information & Optimization Sciences*, 17, 243-253.
- A. W. Tucker (1956), Dual systems of homogeneous linear relations; in H. W. Kuhn and A. W. Tucker (Eds.), *Linear Inequalities and Related Systems*, Princeton University Press, Princeton, 3-18.